

PCT

ORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE
Bureau international



DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

<p>(51) Classification internationale des brevets ⁶ : C12N 15/62, 15/14, 15/12, 15/60, 15/80; 15/81, 1/19, 1/15, A61K 38/38, 38/17, C07K 14/765 // C12N 9/88, 1/21, 5/10</p>	<p>A1</p>	<p>(11) Numéro de publication internationale: WO 95/30759 (43) Date de publication internationale: 16 novembre 1995 (16.11.95)</p>
<p>(21) Numéro de la demande internationale: PCT/FR95/00520 (22) Date de dépôt international: 20 avril 1995 (20.04.95) (30) Données relatives à la priorité: 94/05616 6 mai 1994 (06.05.94) FR (71) Déposant (pour tous les Etats désignés sauf US): RHONE-POULENC RORER S.A. [FR/FR]; 20, avenue Raymond-Aron, F-92160 Antony (FR). (72) Inventeurs; et (75) Inventeurs/Déposants (US seulement): BECQUART, Jérôme [FR/FR]; 8, rue Maublanc, F-75015 Paris (FR). CON-SEILLER, Emmanuel [FR/FR]; 10, rue de Plélo, F-75015 Paris (FR). GUITTON, Jean-Dominique [FR/FR]; 74, rue Dunois, F-75013 Paris (FR). HARDY, Florence [FR/FR]; 6, square Villaret-de-Joyeuse, F-75017 Paris (FR). YEH, Patrice [FR/FR]; 11 bis, rue Lacépède, F-75005 Paris (FR). (74) Mandataire: LE COUPANEC, Pascale; Rhône-Poulenc Rorer S.A., Direction Brevets, 20, avenue Raymond-Aron, F-92160 Antony (FR).</p>		<p>(81) Etats désignés: CA, JP, US, brevet européen (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Publiée Avec rapport de recherche internationale.</p>
<p>(54) Title: BIOLOGICALLY ACTIVE POLYPEPTIDES INSERTED INTO AN ALBUMIN (54) Titre: POLYPEPTIDES BIOLOGIQUEMENT ACTIFS INSERES DANS UNE ALBUMINE (57) Abstract Biologically active recombinant polypeptides essentially consisting of at least one active portion derived from a natural or artificial biologically active polypeptide and inserted into an albumin or albumin variant, the preparation thereof, and pharmaceutical compositions containing same, are disclosed. (57) Abrégé La présente invention concerne des polypeptides recombinants biologiquement actifs essentiellement composés d'au moins une partie active dérivée d'un polypeptide, naturel ou artificiel, ayant une activité biologique, insérée dans une albumine ou un variant d'albumine, leur préparation et des compositions pharmaceutiques les contenant.</p>		

BEST AVAILABLE COPY

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AT	Autriche	GB	Royaume-Uni	MR	Mauritanie
AU	Australie	GE	Géorgie	MW	Malawi
BB	Barbade	GN	Guinée	NE	Niger
BE	Belgique	GR	Grèce	NL	Pays-Bas
BF	Burkina Faso	HU	Hongrie	NO	Norvège
BG	Bulgarie	IE	Irlande	NZ	Nouvelle-Zélande
BJ	Bénin	IT	Italie	PL	Pologne
BR	Brésil	JP	Japon	PT	Portugal
BY	Bélarus	KE	Kenya	RO	Roumanie
CA	Canada	KG	Kirghizistan	RU	Fédération de Russie
CF	République centrafricaine	KP	République populaire démocratique de Corée	SD	Soudan
CG	Congo	KR	République de Corée	SE	Suède
CH	Suisse	KZ	Kazakhstan	SI	Slovénie
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovaquie
CM	Cameroun	LK	Sri Lanka	SN	Sénégal
CN	Chine	LU	Luxembourg	TD	Tchad
CS	Tchécoslovaquie	LV	Lettonie	TG	Togo
CZ	République tchèque	MC	Monaco	TJ	Tadjikistan
DE	Allemagne	MD	République de Moldova	TT	Trinité-et-Tobago
DK	Danemark	MG	Madagascar	UA	Ukraine
ES	Espagne	ML	Mali	US	Etats-Unis d'Amérique
FI	Finlande	MN	Mongolie	UZ	Ouzbékistan
FR	France			VN	Viet Nam
GA	Gabon				

Polypeptides biologiquement actifs insérés dans une albumine

La présente invention concerne de nouveaux polypeptides biologiquement
5 actifs, leur préparation et des compositions pharmaceutiques les contenant.

Quoique possédant une ou plusieurs activités thérapeutiques potentielles, de nombreux polypeptides ne peuvent, malheureusement, pas être exploités pharmaceutiquement. Ceci peut avoir différentes raisons, telles que notamment leur faible stabilité in vivo, leur structure complexe ou fragile, la difficulté de les produire à
10 une échelle industriellement acceptable, etc... De même, certains polypeptides ne donnent pas les résultats attendus in vivo en raison de problèmes d'administration, de conditionnement, de pharmacocinétique etc...

La présente invention a précisément pour objet de remédier à ces inconvénients.

15 Elle vise notamment l'élaboration de protéines artificielles biologiquement actives permettant une exploitation optimale, sur le plan thérapeutique, des propriétés biologiques de ces polypeptides.

La Demanderesse a ainsi mis en évidence qu'il est possible d'insérer par voie génétique toute structure active, dérivée d'un polypeptide biologiquement actif, dans
20 une autre structure protéique constituée d'albumine, sans en altérer lesdites propriétés biologiques. De manière inattendue, la sérum-albumine humaine permet de présenter efficacement la structure active à ses sites d'interaction et d'assurer une stabilité plasmatique élevée au polypeptide recombinant de l'invention.

Plus précisément, la présente invention concerne un polypeptide
25 recombinant comportant au moins une partie active dérivée d'un polypeptide, naturel ou artificiel, biologiquement actif, génétiquement insérée dans une albumine ou un de ses variants ou dérivés.

Par variant de l'albumine, on entend désigner selon la présente invention toute protéine à haute demi-vie plasmatique obtenue par modification, à l'aide des
30 techniques du génie génétique, d'un gène codant pour un isomorphe donné de la

sérum-albumine humaine, ainsi que toute macromolécule à haute demie-vie plasmatique obtenue par modification in vitro de la protéine codée par de tels gènes. (Par modification, on doit entendre toute mutation, substitution, délétion, addition ou modification de nature génétique et/ou chimique). L'albumine étant très polymorphe,
5 de nombreux variants naturels ont été identifiés et répertoriés [Weitkamp L.R. et al., Ann. Hum. Genet. 37 (1973) 219].

En ce qui concerne les dérivés d'albumine, il s'agit plus particulièrement de molécules comportant tout ou partie de l'albumine, fusionnée le cas échéant à au moins une séquence polypeptidique provenant d'un gène naturel ou artificiel, elle
10 même dotée ou non d'une activité biologique.

Dans la suite de la description, les différents types d'albumines explicités ci-dessus sont communément désignés sous le terme albumine.

Au sens de la présente invention, il est entendu par partie active, une partie possédant une activité qui peut soit être directe (traitement des maladies, diagnostic,
15 recherche biologique, capteurs...), ou indirecte (par exemple utilisable dans la prévention des maladies, dans la conception des vaccins, dans les techniques de l'imagerie médicale etc...).

Les parties actives de polypeptides biologiquement actifs, insérées selon l'invention, présentent de préférence un intérêt thérapeutique.

20 Les polypeptides possédant une activité thérapeutique peuvent être d'origine humaine ou non.

A titre représentatif des polypeptides d'origine non humaine, on peut citer des peptides ou leurs dérivés, possédant des propriétés potentiellement utiles dans les pathologies des compartiments sanguins et interstitiels, tels que l'hirudine, la
25 trigramine, l'antistatine, les peptides anticoagulant des tiques (TAP), l'ariétine, l'aplagine etc....

Selon un mode privilégié de l'invention, le polypeptide ayant une activité thérapeutique est un polypeptide d'origine humaine ou un variant moléculaire. Par exemple, il peut s'agir de tout ou partie, d'un enzyme, d'un inhibiteur d'enzyme, d'un
30 antigène, d'un anticorps, d'une hormone, d'un récepteur, d'un facteur intervenant

dans le contrôle de la coagulation, d'un interféron, d'une cytokine [les interleukines, mais aussi leurs variants antagonistes naturels de leur fixation au(x) récepteur(s), les cytokines de type SIS (small induced secreted) et par exemple les protéines inflammatoires des macrophages (les MIPs), etc...], d'un facteur de croissance et/ou de différenciation [et par exemple les facteurs de croissance transformants (les TGFs), les facteurs de différenciation des cellules sanguines (érythropoïétine, M-CSF, G-CSF, GM-CSF etc...), l'insuline et les facteurs de croissance qui lui ressemblent (les IGFs), ou encore les facteurs de perméabilité cellulaire (VPF/VEGF), etc...], d'un facteur impliqué dans la génèse/résorption des tissus osseux (OIF et ostéopontine par exemple), d'un facteur impliqué dans la motilité ou la migration cellulaire [et par exemple le facteur de mobilité autocrine (AMF), le facteur de stimulation de la migration (MSF), ou encore le facteur de dispersion (scatter factor/facteur de croissance des hépatocytes)], d'un facteur bactéricide ou antifongique, d'un facteur chimiotactique [et par exemple le facteur plaquettaire 4 (PF4), ou encore les peptides chemoattractants des monocytes (MCP/MCAF) ou des neutrophiles (NCAF), etc...], d'un facteur cytostatique (et par exemple les protéines qui se fixent aux galactosides), d'une molécule adhésive plasmatique (et par exemple le facteur de von Willebrand, le fibrinogène etc...) ou interstitielle (laminine, ténascine, vitronectine, etc...) ou des matrices extracellulaires, ou encore toute séquence peptidique antagoniste ou agoniste d'interactions moléculaires et/ou intercellulaires impliquées dans les pathologies des compartiments circulatoires et interstitiels et par exemple la formation des thrombus artériels et veineux, les métastases cancéreuses, l'angiogénèse tumorale, le choc inflammatoire, les maladies autoimmunes, les pathologies osseuses et ostéo-articulaires etc...

Bien entendu, la partie active des polypeptides de l'invention peut être constituée par le polypeptide entier biologiquement actif ou par une structure dérivée de celui-ci ou encore correspondre à une séquence peptidique non naturelle, isolée par exemple à partir de banques peptidiques aléatoires. (Pour des raisons de simplification, ces différentes possibilités seront couvertes dans ce qui suit sous la désignation commune "partie active d'un peptide biologiquement actif".) Au sens de la présente invention, on entend par structure dérivée tout polypeptide obtenu par modification et conservant une activité biologique. Par modification, on doit entendre toute mutation, substitution, délétion, addition ou modification de nature génétique et/ou chimique. De tels dérivés peuvent être générés dans des buts différents, tels

que notamment celui d'augmenter l'affinité de la molécule pour ses sites de fixation, celui d'améliorer ses niveaux de production, celui d'augmenter sa résistances aux protéases, celui d'augmenter son efficacité thérapeutique ou encore de réduire ses effets secondaires, ou celui de lui conférer de nouvelles propriétés biologiques. A titre
5 d'exemple, les polypeptides chimères de l'invention possèdent des propriétés pharmacocinétiques et une activité biologique utilisable pour la prévention ou le traitement des maladies.

Des polypeptides de l'invention, particulièrement avantageux, sont ceux dans lesquels la partie active présente :

- 10 (a) la structure peptidique entière ou,
(b) un fragment de (a) ou une structure dérivée de (a) par modification structurale (mutation, substitution addition et/ou délétion d'un ou plusieurs résidus) et possédant une activité thérapeutique.

Parmi les structures du type (b), on peut citer plus particulièrement les
15 molécules dans lesquelles certains sites de N- ou O-glycosylation ont été modifiés ou supprimés, les molécules dans lesquelles un ou plusieurs résidus ont été substitués, ou les molécules dans lesquelles tous les résidus cystéine ont été substitués. On peut citer également des molécules obtenues à partir de (a) par délétion de régions n'intervenant pas ou peu dans l'interaction avec les sites de liaison considérés, ou
20 exprimant une activité indésirable, et des molécules comportant par rapport à (a) des résidus supplémentaires, tels que par exemple une méthionine N-terminale et/ou un signal de sécrétion et/ou un peptide de jonction.

L'objet de la présente invention est particulièrement avantageux pour les peptides actifs trop petits pour former un domaine structural et/ou ne possédant pas
25 une bonne stabilité in vivo et/ou une bonne biodisponibilité. L'insertion proposée selon l'invention, permet de les associer à un ou des domaine(s) pré-existant(s) de l'albumine et de bénéficier ainsi de la biodisponibilité et de la stabilité in vivo de celle-ci.

De manière générale, la taille des parties actives insérées dans l'albumine
30 varie entre trois à vingt cinq résidus d'acides aminés. Toutefois, des séquences allant de 1 résidu à 100 résidus peuvent également être utilisées.

L'insertion d'une partie active de peptide dans la séquence peptidique de l'albumine est réalisée selon l'invention de manière à satisfaire aux deux conditions suivantes:

Il doit être préservé à ladite partie active, insérée au sein de l'albumine, une
5 accessibilité suffisante afin de lui conserver intacte son activité biologique. Par ailleurs, la structure de l'albumine ne doit pas non plus subir une déstabilisation trop importante qui serait bien entendu préjudiciable au polypeptide recombinant dit la chimère.

Les sites d'insertion sont préférentiellement sélectionnés au sein de
10 l'albumine en respectant les précédents impératifs.

Selon la structure cristalline publiée par He et Carter (Nature 1992, 358, 209-215) l'albumine est formée de la répétition de 3 domaines comprenant chacun deux sous-domaines et elle contient plus de 67% d'hélices alpha. Chacun des domaines est superposable aux autres et est formé de 10 hélices notées de h1 à h10. Le sous-
15 domaine A comporte les hélices h1 à h6 et le sous-domaine B, les hélices h7 à h10. Chaque sous domaine est formé par un motif commun: h1, h2, h3, h4 pour le domaine A et h7, h8, h9 et h10 pour le domaine B. Les petites hélices h5 et h6 supplémentaires sont liées par un pont disulfure au sous domaine A. La figure 1 rend compte de manière schématique de la structure de l'albumine humaine.

20 Les sites d'insertion sont de préférence localisés dans les régions de l'albumine présumées former des régions exposées à la surface de la molécule, ces régions étant préférentiellement des boucles.

A titre de sites d'insertion convenant tout particulièrement à l'invention, on peut mentionner trois régions du premier domaine :

- 25 - la région 5 qui s'étend du résidu 57 à 62 et qui correspond à une boucle reliant les deux hélices h3 et h4 ;
- la région 8 comprenant les résidus 103 à 120 et correspondant à la zone inter sous-domaines.
- la région 13 comprise entre les résidus 178 et 200 et qui correspond à une hélice.

30 Les hélices h2 et h3 du domaine III délimitent également, du résidu 419 au résidu 430, une autre région d'insertion envisageable selon l'invention.

Une partie active de peptide biologiquement actif peut être insérée selon trois modes différents dans la séquence peptidique de l'albumine.

- Il peut s'agir d'une insertion stricte consistant en une simple addition de la séquence du peptide d'intérêt dans la séquence d'origine de l'albumine qui est conservée dans sa totalité.
- L'insertion peut correspondre à une substitution d'une portion de la séquence peptidique de l'albumine par la séquence peptidique correspondant à la partie active du peptide d'intérêt.
- Enfin, il peut s'agir d'une insertion combinant une addition d'une partie de la séquence peptidique active et une substitution d'une portion de la séquence peptidique de l'albumine par le reste de la partie active du peptide actif.

10 La figure 2 rend compte d'une manière schématique de ces différents modes d'insertion.

Bien entendu, la partie active d'un peptide biologiquement actif peut être répétée plusieurs fois dans la chimère au même endroit et/ou dans des régions différentes de l'albumine. De même, il est également possible d'insérer selon
15 l'invention des parties actives différentes, soit issues d'un même peptide ou de peptides différents.

Enfin, une partie active des peptides selon l'invention peut être insérée soit strictement au sein de l'albumine soit entourée de séquences de jonction.

En ce qui concerne ces séquences de jonction, il peut notamment s'agir de
20 séquences peptidiques riches en résidus glycine et/ou en résidus sérine et/ou en résidus thréonine et/ou tout résidu d'acide aminé décrit comme fréquemment rencontré dans les zones de flexibilité dans les protéines.

Les polypeptides recombinants de l'invention s'avèrent tout particulièrement
25 avantageux.

Ils permettent de maintenir dans l'organisme une activité biologique donnée pendant un temps prolongé. Il s'avère ainsi possible de réduire les doses administrées et, dans certains cas, de potentialiser l'effet thérapeutique, par exemple en réduisant les effets secondaires consécutifs à une administration plus importante.
30 Ils permettent avantageusement de générer et d'utiliser des structures dérivées de polypeptides biologiquement actifs très petites et donc très spécifiques d'un effet recherché. Ces polypeptides recombinants possèdent par ailleurs une répartition particulièrement avantageuse dans l'organisme, ce qui modifie leurs propriétés pharmacocinétiques et favorise le développement de leur activité biologique et leur

utilisation. Ils présentent également l'avantage d'être faiblement ou non-immunogéniques pour l'organisme dans lequel ils sont utilisés. Finalement, les polypeptides de l'invention peuvent être exprimés (et préférentiellement sécrétés) par des organismes recombinants, à des niveaux permettant leur exploitation industrielle.

5 Un autre objet de l'invention concerne un procédé de préparation des molécules chimères décrites ci-avant. Plus précisément, ce procédé consiste à faire exprimer par un hôte cellulaire eucaryote ou procaryote une séquence nucléotidique codant pour une partie active d'un polypeptide désiré, puis à récolter le polypeptide produit.

10 Parmi les hôtes eucaryotes utilisables dans le cadre de la présente invention, on peut citer les cellules animales, les levures, ou les champignons. En particulier, s'agissant de levures, on peut citer les levures du genre Saccharomyces, Kluyveromyces, Pichia, Schwanniomyces, ou Hansenula. S'agissant de cellules animales, on peut citer les cellules COS, CHO, Cl27, etc... Parmi les champignons
15 susceptibles d'être utilisés dans la présente invention, on peut citer plus particulièrement Aspergillus ssp. ou Trichoderma ssp. Comme hôtes procaryotes, on préfère utiliser les bactéries telles que Escherichia coli, ou appartenant aux genres Corynebacterium, Bacillus, ou Streptomyces.

20 Les séquences nucléotidiques utilisables dans le cadre de la présente invention peuvent être préparées de différentes manières. Généralement, elles sont obtenues en assemblant en phase de lecture les séquences codant pour chacune des parties fonctionnelles du polypeptide. Celles-ci peuvent être isolées par les techniques de l'homme de l'art, et par exemple directement à partir des ARN messagers (ARNm) cellulaires, ou par reclonage à partir d'une banque d'ADN
25 complémentaire (ADNc), ou encore il peut s'agir de séquences nucléotidiques totalement synthétiques. Il est entendu de plus que les séquences nucléotidiques peuvent également être ultérieurement modifiées, par exemple par les techniques du génie génétique, pour obtenir des dérivés ou des variants desdites séquences.

30 L'insertion de cette séquence nucléotidique, codant pour la partie active du polypeptide, peut être réalisée directement ou non, selon la région choisie pour site d'insertion, dans le gène codant pour l'albumine.

La région sélectionnée peut, en effet, ne pas comporter de site de restriction adéquat à la réalisation de ladite insertion. Dans cette hypothèse, il peut s'avérer intéressant, préalablement à l'insertion, d'introduire un ou plusieurs sites de restriction

uniques. La création de sites de restriction, au niveau de la région choisie pour site d'insertion, se fait de préférence par mutagénèse dirigée selon des techniques classiques. Toutefois on peut également insérer la séquence nucléotidique correspondant à "la partie active du polypeptide biologiquement actif" directement par mutagénèse dirigée sans que l'insertion ne fasse apparaître de sites de restriction particuliers.

Dans le cas particulier de la région 5 du gène codant pour l'albumine, la présence du site de restriction PVuII autorise l'insertion directe de la séquence nucléotidique. Il se révèle donc particulièrement utile comme site de clonage d'une partie active d'un polypeptide que l'on souhaite insérer en phase traductionnelle dans la séquence de l'albumine au niveau du 57^{ème} résidu.

Dans ce mode d'insertion, mettant à profit un site de restriction unique déjà existant dans la séquence d'origine de l'albumine, la ligature de la séquence codant pour le peptide actif avec le fragment de restriction, correspondant à la totalité du gène codant pour l'albumine, génère une séquence nucléotidique comportant un gène hybride codant pour une protéine chimère du type SAH insertion stricte.

Dans le cas particulier de la zone 419 à 430, la présence de 2 sites uniques d'insertion, (HincII et ArvII) permet d'insérer et/ou substituer les peptides d'intérêt biologique à la séquence de l'albumine.

En ce qui concerne plus particulièrement les régions 8 et 13, l'insertion de ladite séquence est favorisée si l'on y crée préalablement un ou des sites de restriction manipulables. Bien entendu, les sites de restriction à créer au niveau de la séquence sont choisis en tenant compte de la nature des sites de restriction déjà existants. Ils doivent conduire à une insertion sélective.

Dans ce second mode de réalisation, l'emploi de deux sites de restriction uniques permet de construire des gènes codant pour des chimères ayant le peptide actif en insertion et/ou en substitution. L'insertion du peptide se fait par remplacement d'un fragment borné de deux sites de restriction uniques, dans l'ADN complémentaire de l'albumine, insérés par mutagénèse dirigée. Ces deux sites de restriction uniques pourront être, Mst I et Kpn I, dans la région 8 et Sst I et Xho I dans la région 13. La création de ces sites de restriction peut ou non modifier la séquence polypeptidique de l'albumine humaine. Le clonage subséquent du peptide actif en phase codante dans ce gène de l'albumine peut être exclusivement la séquence codante du peptide

ou peut correspondre à un panachage de la séquence codante du peptide et de la séquence codante du fragment délété de l'albumine.

La présente invention vise également la protection des variants de
5 séquences nucléotidiques codant pour l'albumine correspondante c'est à dire intégrant au moins un site de restriction unique non naturel c'est à dire non présent dans la séquence d'origine.

Avantageusement, la création de ces sites de restriction peut servir par la
suite à l'insertion d'un ou plusieurs parties actives de polypeptide(s) biologiquement
10 actif(s) dans la protéine mature.

Plus préférentiellement, dans le procédé de l'invention, la séquence
nucléotidique fait partie d'une cassette d'expression comprenant une région d'initiation
de la transcription (région promoteur) permettant, dans les cellules hôtes, l'expression
15 de la séquence nucléotidique placée sous son contrôle et codant pour les
polypeptides de l'invention. Cette région peut provenir de régions promoteurs de
gènes fortement exprimés dans la cellule hôte utilisée, l'expression étant constitutive
ou régulable. S'agissant de levures, il peut s'agir du promoteur du gène de la
phosphoglycérate kinase (PGK), de la glyceraldéhyde-3-phosphate déshydrogénase
20 (GPD), de la lactase (LAC4), des émolases (ENO), des alcools deshydrogénases
(ADH), etc... S'agissant de bactéries, il peut s'agir du promoteur des gènes droit ou
gauche du bactériophage lambda (P_L, P_R), ou encore des promoteurs des gènes des
opérons tryptophane (P_{trp}) ou lactose (P_{lac}). En outre, cette région de contrôle peut
être modifiée, par exemple par mutagénèse in vitro, par introduction d'éléments
25 additionnels de contrôle ou de séquences synthétiques, ou par des délétions ou des
substitutions des éléments originels de contrôle. La cassette d'expression peut
également comprendre une région de terminaison de la transcription fonctionnelle
dans l'hôte envisagé, positionnée immédiatement en aval de la séquence
nucléotidique codant pour un polypeptide de l'invention.

30 Dans un mode préféré, les polypeptides de l'invention résultent de
l'expression dans un hôte eucaryote ou procaryote d'une séquence nucléotidique et
de la sécrétion du produit d'expression de ladite séquence dans le milieu de culture. Il
est en effet particulièrement avantageux de pouvoir obtenir par voie recombinante
des molécules directement dans le milieu de culture. Dans ce cas, la séquence

nucléotidique codant pour un polypeptide de l'invention est précédée d'une séquence "leader" (ou séquence signal) dirigeant le polypeptide naissant dans les voies de sécrétion de l'hôte utilisé. Cette séquence "leader" peut être la séquence signal naturelle du polypeptide biologiquement actif dans le cas où celui-ci est une protéine
5 naturellement sécrétée, ou celle de la structure stabilisatrice, mais il peut également s'agir de toute autre séquence "leader" fonctionnelle, ou d'une séquence "leader" artificielle. Le choix de l'une ou l'autre de ces séquences est notamment guidé par l'hôte utilisé. Des exemples de séquences signal fonctionnelles incluent celles des gènes des phéromones sexuelles ou des toxines "killer" de levures.

10 En plus de la cassette d'expression, un ou plusieurs marqueurs permettant de sélectionner l'hôte recombiné peuvent être additionnés, tels que par exemple le gène URA3 de la levure S. cerevisiae, ou des gènes conférant la résistance à des antibiotiques comme la généticine (G418) ou à tout autre composé toxique comme certains ions métalliques.

15 L'ensemble constitué par la cassette d'expression et par le marqueur de sélection peut être introduit directement dans les cellules hôtes considérées, soit inséré préalablement dans un vecteur autorépliquatif fonctionnel. Dans le premier cas, des séquences homologues à des régions présentes dans le génôme des cellules hôtes sont préférentiellement additionnées à cet ensemble; lesdites séquences étant
20 alors positionnées de chaque côté de la cassette d'expression et du gène de sélection de façon à augmenter la fréquence d'intégration de l'ensemble dans le génôme de l'hôte en ciblant l'intégration des séquences par recombinaison homologue. Dans le cas où la cassette d'expression est insérée dans un système répliquatif, un système de réplication préféré pour les levures du genre Kluyveromyces
25 est dérivé du plasmide pKD1 initialement isolé de K. drosophilum; un système préféré de réplication pour les levures du genre Saccharomyces est dérivé du plasmide 2 μ de S. cerevisiae. De plus, ce plasmide d'expression peut contenir tout ou partie desdits systèmes de réplication, ou peut combiner des éléments dérivés du plasmide pKD1 aussi bien que du plasmide 2 μ .

30 En outre, les plasmides d'expression peuvent être des vecteurs navettes entre un hôte bactérien tel que Escherichia coli et la cellule hôte choisie. Dans ce cas, une origine de réplication et un marqueur de sélection fonctionnant dans l'hôte bactérien sont requises. Il est également possible de positionner des sites de

restriction entourant les séquences bactériennes et uniques sur le vecteur d'expression: ceci permet de supprimer ces séquences par coupure et religature in vitro du vecteur tronqué avant transformation des cellules hôtes, ce qui peut résulter en une augmentation du nombre de copies et en une stabilité accrue des plasmides d'expression dans lesdits hôtes. Par exemple, de tels sites de restriction peuvent correspondre aux séquences telles que 5'-GGCCNNNNNGGCC-3' (SfiI) (SEQ ID N°1) ou 5'-GCGGCCGC-3' (NotI) (SEQ ID N°2) dans la mesure où ces sites sont extrêmement rares et généralement absents d'un vecteur d'expression.

Après construction de tels vecteurs ou cassette d'expression, ceux-ci sont introduits dans les cellules hôtes retenues selon les techniques classiques décrites dans la littérature. A cet égard, toute méthode permettant d'introduire un ADN étranger dans une cellule peut être utilisée. Il peut s'agir notamment de transformation, électroporation, conjugaison, ou toute autre technique connue de l'homme de l'art. A titre d'exemple pour les hôtes de type levure, les différentes souches de Kluyveromyces utilisées ont été transformées en traitant les cellules entières en présence d'acétate de lithium et de polyéthylène glycol, selon la technique décrite par Ito et al. [J. Bacteriol. 153 (1983) 163]. La technique de transformation décrite par Durrens et al. [Curr. Genet. 18 (1990) 7] utilisant l'éthylène glycol et le diméthylsulfoxyde a également été utilisée. Il est aussi possible de transformer les levures par électroporation, selon la méthode décrite par Karube et al. [FEBS Letters 182 (1985) 90]. Un protocole alternatif est également décrit en détail dans les exemples qui suivent.

Après sélection des cellules transformées, les cellules exprimant lesdits polypeptides sont inoculées et la récupération desdits polypeptides peut être faite, soit au cours de la croissance cellulaire pour les procédés "en continu", soit en fin de croissance pour les cultures "en lots" ("batch"). Les polypeptides qui font l'objet de la présente invention sont ensuite purifiés à partir du surnageant de culture en vue de leur caractérisation moléculaire, pharmacocinétique et biologique.

Un système d'expression préféré des polypeptides de l'invention consiste en l'utilisation des levures du genre Kluyveromyces comme cellule hôte, transformées par certains vecteurs dérivés du réplicon extrachromosomique pKD1 initialement isolé chez K. marxianus var. drosophilum. Ces levures, et en particulier K. lactis et K. fragilis sont généralement capables de répliquer lesdits vecteurs de façon stable et possèdent en outre l'avantage d'être incluses dans la liste des organismes G.R.A.S.

("Generally Recognized As Safe"). Des levures privilégiées sont préférentiellement des souches industrielles du genre Kluyveromyces capables de répliquer de façon stable lesdits plasmides dérivés du plasmide pKD1 et dans lesquels a été inséré un marqueur de sélection ainsi qu'une cassette d'expression permettant la sécrétion à des niveaux élevés des polypeptides de l'invention.

La présente invention concerne également les séquences nucléotidiques codant pour les polypeptides chimères décrits ci-avant, ainsi que les cellules recombinantes, eucaryotes ou procaryotes, comprenant de telles séquences.

La présente invention concerne aussi l'application à titre de médicament des polypeptides selon la présente invention. Plus particulièrement, l'invention a pour objet toute composition pharmaceutique comprenant un ou plusieurs polypeptides ou séquences nucléotidiques tels que décrits ci-avant. Ces compositions pharmaceutiques peuvent se présenter sous forme de diverses formulations. Il peut notamment s'agir de nanoparticules à la surface desquelles sont présents des polypeptides selon l'invention. Ce type de formulation est plus particulièrement utilisé pour effectuer des ciblage dirigés en principe actif. Bien entendu, les séquences nucléotidiques peuvent être utilisées en thérapie génique.

La présente invention sera plus complètement décrite à l'aide des exemples qui suivent, qui doivent être considérés comme illustratifs et non limitatifs.

LISTE DES FIGURES

Les représentations des plasmides indiquées dans les Figures suivantes ne sont pas tracées à l'échelle et seuls les sites de restriction importants pour la compréhension des clonages réalisés ont été indiqués.

Figure 1: Représentation schématique du domaine I de l'albumine avec localisation de sites d'insertion selon l'invention. * signale la localisation du site d'insertion h2-h3 dans le domaine III. Les chiffres 5, 8 et 13 identifient les zones d'insertion correspondantes.

Figure 2 : Représentation schématique de différents modes d'insertion d'un peptide actif dans la structure de l'albumine.

Figure 3: Plasmide pYG105.

Figure 4: Modification de la région 5 de l'albumine suite à l'insertion de la séquence codant pour IEGR telle que décrite dans l'exemple 8.1. Les modifications figurent en caractères gras. L'emplacement des sites de restriction modifiés ou apportés sont indiqués par un trait horizontal et la position de coupure des enzymes correspondant par un trait vertical.

Figure 5: Stratégie de clonage pour l'insertion de IEGR dans la région 5 (exemple 8.1).

Figure 6: Stratégie de clonage pour l'insertion d'un peptide actif dans la région 13 de l'albumine.

EXEMPLES

TECHNIQUES GENERALES DE CLONAGE

Les méthodes classiquement utilisées en biologie moléculaire telles que les extractions préparatives d'ADN plasmidique, la centrifugation d'ADN plasmidique en gradient de chlorure de césium, l'électrophorèse sur gels d'agarose ou d'acrylamide, la purification de fragments d'ADN par électroélution, les extraction de protéines au phénol ou au phénol-chloroforme, la précipitation d'ADN en milieu salin par de l'éthanol ou de l'isopropanol, la transformation dans *Escherichia coli* etc ... sont bien connues de l'homme de métier et sont abondamment décrites dans la littérature [Maniatis T. et al., "Molecular Cloning, a Laboratory Manual", Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982 ; Ausubel F.M. et al. (eds), "Current Protocols in Molecular Biology", John Wiley & Sons, New York, 1987].

Les enzymes de restriction ont été fournies par New England Biolabs (Biolabs), Bethesda Research Laboratories (BRL) ou Amersham et sont utilisées selon les recommandations des fournisseurs.

Les plasmides de type pBR322, pUC et les phages de la série M13 sont d'origine commerciale (Bethesda Research Laboratories).

Pour les ligatures, les fragments d'ADN sont séparés selon leur taille par électrophorèse en gels d'agarose ou d'acrylamide, extraits au phénol ou par un

mélange phénol/chloroforme, précipités à l'éthanol puis incubés en présence de l'ADN ligase du phage T4 (Biolabs) selon les recommandations du fournisseur.

Le remplissage des extrémités 5' proéminentes est effectué par le fragment de Klenow de l'ADN Polymérase I d'E. coli (Biolabs) selon les spécifications du fournisseur. La destruction des extrémités 3' proéminentes est effectuée en présence de l'ADN Polymérase du phage T4 (Biolabs) utilisée selon les recommandations du fabricant. La destruction des extrémités 5' proéminentes est effectuée par un traitement ménagé par la nucléase S1.

La mutagénèse dirigée in vitro par oligodéoxynucléotides synthétiques est effectuée selon la méthode développée par Taylor et al. [Nucleic Acids Res. 13 (1985) 8749-8764] en utilisant le kit distribué par Amersham. Cette technique est notamment mise en oeuvre, dans le cadre de la présente invention, pour créer des sites de restriction uniques en vu d'une insertion subséquente.

L'amplification enzymatique de fragments d'ADN par la technique dite de PCR [Polymérase-catalyzed Chain Reaction, Saiki R.K. et al., Science 230 (1985) 1350-1354; Mullis K.B. et Faloona F.A., Meth. Enzym. 155 (1987) 335-350] est effectuée en utilisant un "DNA thermal cycler" (Perkin Elmer Cetus) selon les spécifications du fabricant.

La vérification des séquences nucléotidiques est effectuée par la méthode développée par Sanger et al. [Proc. Natl. Acad. Sci. USA, 74 (1977) 5463-5467] en utilisant le kit distribué par Amersham.

Les transformations de K. lactis avec l'ADN des plasmides d'expression des protéines de la présente invention sont effectuées par toute technique connue de l'homme de l'art, et dont un exemple est donné dans le texte.

Sauf indication contraire, les souches bactériennes utilisées sont E. coli MC1060 (lacIPOZYA, X74, galU, galK, strA^r), ou E. coli TG1 (lac, proA,B, supE, thi, hsdD5 / F'traD36, proA⁺B⁺, lacI^q, lacZ, M15).

Les souches de levures utilisées appartiennent aux levures bourgeonnantes et plus particulièrement aux levures du genre Kluyveromyces. Les souche K. lactis MW98-8C (a, uraA, arg, lys, K⁺, pKD1^o) et K. lactis CBS 293.91 ont été particulièrement utilisées; un échantillon de la souche MW98-8C a été déposé le 16 Septembre 1988 au Centraalbureau voor Schimmelkulturen (CBS) à Baarn (Pays Bas) où il a été enregistré sous le numéro CBS 579.88.

Une souche bactérienne (E. coli) transformée avec le plasmide pET-8c52K a été déposée le 17 Avril 1990 auprès de l'American Type Culture Collection sous le numéro ATCC 68306.

Les souches de levures transformées par les plasmides d'expression codant pour les protéines de la présente invention sont cultivées en erlenmeyers ou en fermenteurs pilotes de 2l (SETRIC, France) à 28°C en milieu riche (YPD: 1% yeast extract, 2% Bactopeptone, 2% glucose; ou YPL: 1% yeast extract, 2% Bactopeptone, 2% lactose) sous agitation constante.

10

EXEMPLE 1:

PROTOCOLE D'INSERTION STRICTE DE PEPTIDE EN UTILISANT UN SITE DE RESTRICTION UNIQUE, PRESENT DANS LE GENE DE LA SAH.

De par son unicité dans le gène HSA et le vecteur associé, le site PvuII localisé naturellement dans la séquence codante, est particulièrement utile comme site de clonage d'un peptide biologiquement actif que l'on désire insérer en phase traductionnelle dans la SAH au niveau du 58^{ième} résidu. Dans un mode de réalisation particulier, il est utile d'employer des peptides de p résidus dont la séquence codante est [3xN]_p. Dans ce cas, les oligonucléotides synthétisés sont du type 5'-NN [3xN]_pN-3' et son brin complémentaire. La ligature de ce fragment avec le fragment de restriction HindIII-HindIII Δ PvuII, correspondant à la totalité du gène codant pour la SAH, génère un fragment de restriction HindIII-HindIII comportant un gène hybride codant pour une protéine chimère du type SAH insertion stricte. Dans un autre mode de réalisation, le peptide peut être répété plusieurs fois dans la chimère.

EXEMPLE 2:

PROTOCOLE D'INSERTION TOTALE OU PARTIELLE DE PEPTIDE EN UTILISANT DEUX SITES DE RESTRICTION UNIQUES ET NATURELS.

Dans un mode de réalisation particulier, l'utilisation de deux sites de restriction uniques permet de construire des gènes codant pour des chimères ayant le peptide actif en insertion stricte ou en substitution ou en insertion partielle.

Par exemple, l'existence des sites uniques HincII et AvrII dans la séquence codante de SAH et dans le vecteur permet de générer un fragment HindIII-HindIII Δ HincII-AvrII. L'élimination du fragment nucléotidique HincII-AvrII correspond à la délétion du fragment peptidique T(420) - N(429). L'utilisation d'un oligonucléotide
5 complémentaire approprié permet de cloner un peptide en phase codante dans le gène de la SAH. Ce fragment de restriction peut être exclusivement la séquence codante complémentaire du peptide ou peut correspondre à un panachage de la séquence codante du peptide actif et de la séquence codante du fragment T(420) - N(429) de la SAH. Le peptide actif peut être présent plusieurs fois dans la chimère.

10

EXEMPLE 3

PLASMIDES D'EXPRESSION

Les protéines chimères des exemples précédents peuvent être exprimées dans les levures à partir de promoteurs fonctionnels, régulables ou constitutifs, tels que, par exemple, ceux présents dans les plasmides pYG105 (promoteur LAC4 de Kluyveromyces lactis), pYG106 (promoteur PGK de Saccharomyces cerevisiae), pYG536 (promoteur PHO5 de S.cerevisiae), ou des promoteur hybrides tels que ceux décrits dans la demande de brevet EP 361 991. Les plasmides pYG105 et pYG106 sont ici particulièrement utiles car ils permettent l'expression des gènes codés par les
15 fragments de restriction HindIII tel que décrits dans les exemples précédents et clonés dans le site HindIII et dans l'orientation productive (définie comme l'orientation qui place la région "prépro" de l'albumine de façon proximale par rapport au promoteur de transcription), à partir de promoteurs fonctionnels chez K.lactis, régulables (pYG105) ou constitutifs (pYG106). Le plasmide pYG105 correspond au plasmide pKan707
20 décrit dans la demande de brevet EP 361 991 dans lequel le site de restriction HindIII unique et localisé dans le gène de résistance à la généticine (G418) a été détruit par mutagenèse dirigée tout en conservant une protéine inchangée (oligodeoxynucleotide 5'-GAAATGCATAAGCTCTTGCCATTCTACCG-3' (SEQ ID N°3). Le fragment Sall-SacI codant pour le gène URA3 du plasmide muté a été ensuite remplacé par un
25 fragment de restriction Sall-SacI comportant une cassette d'expression constituée du promoteur LAC4 de K. lactis (sous la forme d'un fragment Sall-HindIII) et du terminateur du gène PGK de S. cerevisiae (sous la forme d'un fragment HindIII-SacI).
30 Le plasmide pYG105 est mitotiquement très stable chez les levures Kluyveromyces. Il

est représenté en figure 3. Les plasmides pYG105 et pYG106 ne diffèrent entre eux que par la nature du promoteur de transcription encodé par le fragment Sall-HindIII.

EXEMPLE 4 :**TRANSFORMATION DES LEVURES**

La transformation des levures appartenant au genre *Kluyveromyces*, et en particulier les souches MW98-8C et CBS 293.91 de *K. lactis*, s'effectue par exemple par la technique de traitement des cellules entières par de l'acétate de lithium [Ito H. et al., J. Bacteriol. 153 (1983) 163-168], adaptée comme suit. La croissance des cellules se fait à 28°C dans 50 ml de milieu YPD, avec agitation et jusqu'à une densité optique à 600 nm (DO_{600}) comprise entre 0,6 et 0,8; les cellules sont récoltées par centrifugation à faible vitesse, lavées dans une solution stérile de TE (10 mM Tris HCl pH 7,4; 1 mM EDTA), resuspendues dans 3-4 ml d'acétate lithium (0,1 M dans du TE) pour obtenir une densité cellulaire d'environ 2×10^8 cellules/ml, puis incubées à 30°C pendant 1 heure sous agitation modérée. Des aliquotes de 0,1 ml de la suspension résultante de cellules compétentes sont incubés à 30°C pendant 1 heure en présence d'ADN et à une concentration finale de 35% de polyéthylène glycol (PEG4000, Sigma). Après un choc thermique de 5 minutes à 42°C, les cellules sont lavées 2 fois, resuspendues dans 0,2 ml d'eau stérile et incubées 16 heures à 28°C dans 2 ml de milieu YPD pour permettre l'expression phénotypique du gène de résistance au G418 exprimé sous contrôle du promoteur P_{K1} (cf. EP 361 991); 200 µl de la suspension cellulaire sont ensuite étalés sur boîtes YPD sélectives (G418, 200 µg/ml). Les boîtes sont mises à incuber à 28°C et les transformants apparaissent après 2 à 3 jours de croissance cellulaire.

EXEMPLE 5:**SECRETION DES CHIMERES**

Après sélection sur milieu riche supplémenté en G418 les clones recombinants sont testés pour leur capacité à sécréter la forme mature des protéines chimères. Quelques clones correspondant à la souche CBS 293.91 ou MW98-8C transformée par les plasmides d'expression des chimères entre la SAH et la partie biologiquement active sont mis à incuber en milieu YPD ou YPL à 28°C. Les surnageants cellulaires sont récupérés par centrifugation quand les cellules atteignent la phase stationnaire de croissance, éventuellement concentrés 10 fois par précipitation pendant 30 minutes à -20°C dans une concentration finale de 60% d'éthanol, puis testés après électrophorèse en gel SDS-PAGE (), soit directement par coloration du gel par du

bleu de coomassie, soit après immunoblot en utilisant des anticorps primaires dirigés contre la partie biologiquement active ou un sérum polyclonal de lapin dirigé contre la SAH. Lors des expériences de détection immunologique, le filtre de nitrocellulose est d'abord incubé en présence des anticorps primaires spécifiques, lavé plusieurs fois, 5 incubé en présence d'anticorps de chèvre dirigés contre les anticorps primaires, puis incubé en présence d'un complexe avidine-péroxydase en utilisant le "kit ABC" distribué par Vectastain (Biosys S.A., Compiègne, France). La réaction immunologique est ensuite révélée par addition de diamino-3,3' benzidine tetrachlorhydrate (Prolabo) en présence d'eau oxygénée, selon les recommandations 10 du fabricant.

EXEMPLE 6: **PURIFICATION DES CHIMERES**

15 Les chimères présentes dans les surnageants de culture correspondant à la souche CBS 293.91 transformée, sont caractérisées dans un premier temps à l'aide d'anticorps spécifiques de la partie SAH. Il peut être souhaitable de purifier certaines de ces chimères. La culture est alors centrifugée (10000g, 30 min), le surnageant est 20 passé à travers un filtre de 0,22 mm (Millipore), puis concentré par ultrafiltration (Amicon) en utilisant une membrane dont le seuil de discrimination se situe à 30 kDa. Le concentrat obtenu est alors dialysé contre une solution de Tris HCl (50 mM pH 8) puis purifié sur colonne. Par exemple, le concentrat correspondant au surnageant de culture de la souche CBS 293.91 transformée est purifié par chromatographie 25 d'affinité sur Bleu-Trisacryl (IBF) et les échantillons sont ensuite dialysés contre de l'eau. Une purification par tamis moléculaire peut être ensuite réalisée. Dans ce cas, une colonne Superose 12 (Pharmacia) est préalablement équilibrée dans du tampon 20mM NaH₂PO₄ 100mM NaCl pH 7.0 et les échantillons issus de la chromatographie d'affinité sont chargés sur la colonne, récoltés et caractérisés. Une purification par 30 chromatographie d'échange d'ions peut également être utilisée. Dans certains cas, le concentrat obtenu après ultrafiltration est dialysé contre une solution de Tris HCl (50 mM pH 8), puis déposé par fractions de 20 ml sur une colonne (5 ml) échangeuse de cations (S Fast Flow, Pharmacia) équilibrée dans le même tampon. La colonne est alors lavée plusieurs fois par la solution de Tris HCl (50 mM pH 8) et la protéine

chimère est alors éluée de la colonne par un gradient (0 à 1 M) de NaCl. Les fractions contenant la protéine chimère sont alors réunies et dialysées contre une solution de Tris HCl 50 mM (pH 8) puis redéposées sur colonne S Fast Flow. Après élution de la colonne, les fractions contenant la protéine sont réunies, dialysées
 5 contre de l'eau et lyophilisées avant caractérisation.

EXEMPLE 7:

INSERTION DU PEPTIDE 11 DANS LA SEQUENCE DE L'ALBUMINE

10

Le peptide 11 a été décrit en tant qu'épitope de la tryptophane synthase (Larvor et al. Mol. Immunol. (1991) 28, 523-531).

1. Insertion stricte dans une région ayant naturellement des sites de restriction 15 uniques

Le peptide 11 a la séquence peptidique suivante dans le sens N a C terminal: HGRVGIYFGMK (SEQ ID N°20). La stratégie d'insertion dans la zone HincII-AvrII a consisté à faire une ligature entre un fragment nucléotidique synthétisé codant pour le
 20 peptide 11 et tel que le cadre de lecture pour la séquence codante de la SAH soit respecté et que la nature de la séquence codante de la SAH soit celle d'origine. Dans un premier temps nous avons donc synthétisé les deux oligonucléotides suivants : 5'-CCATGGTAGAGTAGGTATCTATTTCTGGTATGAAAACCTCCAACCTCTTGTAGAGGTCT
 CGAGAAAT-3' (SEQ ID N°4) et 5'-
 25 CTAGATTTCTCGAGACCTCTACAAGATGTGGAGTTTTCATACCGAAATAGATACCT
 ACTCTACCATGG-3' (SEQ ID N°5). Ces deux oligonucléotides ont été hybridés puis ligaturés au fragment HindIII-HindIII Δ HincII-AvrII, générant ainsi la totalité du gène de la SAH comportant l'insertion totale du peptide 11 entre S(419) et T(420), immédiatement précédée de la région d'exportation "prépro" de la SAH. Ce fragment est
 30 cloné dans l'orientation productive et dans le site HindIII du plasmide pYG105.

2. Insertion par substitution et addition dans une région ayant naturellement des sites de restriction uniques

Dans un autre mode de réalisation, les oligonucléotides synthétisés ont été les
 5 suivants : 5'-CCATGGTAGAGTAGGTATCTATTTTCGGTATGAAA-3'
 (SEQ ID N°6) et 5'-CTAGTTTCATACCGAAATAGATACCTACTTCTTACCATGG-3'
 (SEQ ID N°7). Ces deux oligonucléotides sont hybridés puis ligaturés au fragment
HindIII-HindIII Δ HincII-AvrII, générant ainsi le gène d'une chimère amputé des résidus
 T(420) à N(429) et substitués par la séquence du peptide 11. Ce fragment est cloné
 10 dans l'orientation productive et dans le site HindIII du plasmide pYG105.

3. Insertion stricte dans une région ne possédant pas de sites de restriction uniques

Dans un autre mode de réalisation, deux sites de restriction uniques, Sst I et
 15 Xho 1, ont été créés par mutagenèse dirigée. Le même type de clonage directionnel
 a été réalisé. Dans le cas de l'insertion totale du peptide 11 entre les résidus A(191)
 et S (192), les deux oligonucléotides suivants ont été synthétisés :5'-
 ACGGGATGAAGGAAGGCCCATGGTAGAGTAGGTATCTATTTTCGGTATGAAA-3'
 20 (SEQ ID N°8) et 5'-
 TCGATTTTCATTACCGAAATAGATACCTACTCTACCATGGGCCTTCCCTTCATCCCG
 TAGCT-3' (SEQ ID N°9). On procède ensuite, selon le protocole déjà décrit aux
 points 1 ou 2 précédents, pour la réalisation du plasmide d'expression correspondant.

25 4. Insertion par substitution dans une région ne possédant pas de sites de restriction uniques

L'insertion nécessite au préalable la création des deux sites de restriction Sst I
 et Xho 1. Pour l'insertion partielle en soi, les deux oligonucléotides suivants ont été
 30 synthétisés : 5'-ACATGGTAGAGTAGGTATCTATTTTCGGTATGAAA-3' (SEQ ID N°10)
 et 5'-TCGATTTTCATACCGAAATAGATACCTACTCTACCATGTAGCT-3' (SEQ ID
 N°11). Dans ce dernier cas, les résidus R(186) à A(191) sont délétés et substitués
 par le peptide 11. Les deux plasmides d'expression de ces chimères sont
 respectivement 1671 et 1667.

Chacun des plasmides, obtenus aux points 1, 2, 3 et 4, est utilisé pour transformer une souche de levure selon le protocole décrit en exemple 4. les protéines correspondantes sont secrétées et purifiées selon les exemples 5 et 6.

5

EXEMPLE 8**INSERTION DE LA SEQUENCE PEPTIDIQUE IEGR, SUBSTRAT DU FACTEUR Xa**

On procède à l'insertion du peptide contenant la séquence IEGR (SEQ ID N°21), cible du facteur protéasique Xa, qui coupe la prothrombine en thrombine dans la cascade de réactions intervenant dans la coagulation sanguine.

1.Insertion stricte dans la région 5 ayant un site PvuII unique

Pour effectuer cette insertion au site Pvu II du gène de l'albumine, il est dans un premier temps créer un vecteur réplcatif dépourvu du site Pvu II, dans lequel est ensuite inséré le gène codant pour la prépro albumine. Les deux oligonucléotides complémentaires suivants, codant pour la séquence IEGR, ont également été synthétisés: 5'-GATCCATAGAAGGTCGACTAG-3'(SEQ ID N°12) et 3'-CTAGGTATCTTCCAGCTGATC-5'(SEQ ID N°13). Ces deux oligonucléotides sont ensuite hybridés puis insérés au niveau du site Pvu II du gène de l'albumine. Les modifications que leur insertion fait apparaître sur les séquences nucléotidiques et peptidiques de l'albumine sont reportées sur la figure 4. La stratégie de clonage est schématisée en figure 5.

Cette construction illustre le cas où des séquences de jonction sont introduites de part et d'autre du peptide d'intérêt.

2.Insertion dans la région 13.

La figure 6 décrit une stratégie de clonage d'un peptide dans la région 13.

Dans ce qui suit, les sites de restriction Sst I et Xho 1 ont été créés par mutagenèse dirigée. Dans le cas de la région 13 on a procédé respectivement à une substitution totale et une substitution addition de la séquence IEGR simple ou encadrée de séquences de jonction.

En ce qui concerne la substitution totale, les oligonucléotides utilisés sont les suivants:

5'- CAGAATCGAAGGTAGAGCC-3' (SEQ ID N°14) et

5'- TCGAGGCTCTACCTTCGATCGAGGGTAGCT- 3'(SEQ ID N°15)

5 Au niveau protéique, la séquence (187)DEGK (SEQ ID N°22) est substituée par IEGR.

Dans le second cas, les oligonucléotides utilisés sont les suivants:

5'- ACCCTCGATCGAAGGTAGATCTCCA- 3'(SEQ ID N°16)

10 5'- TCGATGGAGATCTACCTTCGATCGAGGGTAGCT-3'(SEQ ID N°17)

Au niveau protéique, l'albumine est amputée du résidu R(186) au résidu A(191) et remplacée par la séquence PSIEGRSP (SEQ ID N°23) conduisant donc à une addition de deux résidus.

15 Les protéines correspondantes sont secrétées et purifiées suivant les protocoles décrits dans les exemples précédents. le tableau I ci-après rend compte des caractéristiques de ces chimères.

Albumine incorporant:	Taux d'expression ug/l	Rendement purification	Pureté finale
IEGR	150	15	98
PSIEGRPS	110	36	96

TABLEAU I

20

3 Activité biologique de la chimère obtenue selon l'exemple 8.1

25 La chimère est incubée en présence de facteur Xa bovin dans un rapport enzyme/substrat de 1/10 et dans un tampon Tris 50mM, NaCl 100mM, CaCl₂ 1mM, pH 8.0 pendant 3 heures à 37°C. A l'issue de ce traitement, on réalise une analyse SDS PAGE qui met en évidence le clivage de la chimère caractérisé par la génération d'un fragment d'albumine de masse de l'ordre de 60kD.

30

LISTE DE SEQUENCES

(1) INFORMATION GENERALE:

- 5 (i) DEPOSANT:
(A) NOM: RHONE-POULENC RORER S.A.
(B) RUE: 20, avenue Raymond ARON
(C) VILLE: ANTONY
(E) PAYS: FRANCE
(F) CODE POSTAL: 92165
- 10 (ii) TITRE DE L' INVENTION :Nouveaux polypeptides biologiquement actifs,
leur preparation et composition pharmaceutique les contenant.
- (iii) NOMBRE DE SEQUENCES: 24
- (iv) FORME LISIBLE PAR ORDINATEUR:
(A) TYPE DE SUPPORT: Tape
15 (B) ORDINATEUR: IBM PC compatible
(C) SYSTEME D' EXPLOITATION: PC-DOS/MS-DOS
(D) LOGICIEL: PatentIn Release #1.0, Version #1.25 (OEB)

(2) INFORMATION POUR LA SEQ ID NO: 1:

- 20 (i) CARACTERISTIQUES DE LA SEQUENCE:
(A) LONGUEUR: 13 paires de base
(B) TYPE: acide nucléique
(C) NOMBRE DE BRINS: simple
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- 25 (iii) HYPOTHETIQUE: NON
(iii) ANTI-SENS: NON
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 1:

GGCCNNNNNG GCC

13

30

(3) INFORMATION POUR LA SEQ ID NO: 2:

- 35 (i) CARACTERISTIQUES DE LA SEQUENCE:
(A) LONGUEUR: 8 paires de base
(B) TYPE: acide nucléique
(C) NOMBRE DE BRINS: simple
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- (iii) HYPOTHETIQUE: NON
(iii) ANTI-SENS: NON
- 40 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 2:

GCGGCCGC

8

(4) INFORMATION POUR LA SEQ ID NO: 3

- 5 (i) CARACTERISTIQUES DE LA SEQUENCE:
(A) LONGUEUR: 30 paires de bases
(B) TYPE: acide nucléique
(C) NOMBRE DE BRINS: simple
(D) CONFIGURATION: linéaire
10 (ii) TYPE DE MOLECULE: ADNc
(iii) HYPOTHETIQUE: NON
(iii) ANTI-SENS: NON
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 3:

GAAATGCATA AGCTCTTGCC ATTCTCACCG

30

15 (5) INFORMATION POUR LA SEQ ID NO: 4:

- 20 (i) CARACTERISTIQUES DE LA SEQUENCE:
(A) LONGUEUR: 64 paires de bases
(B) TYPE: acide nucléique
(C) NOMBRE DE BRINS: simple
(D) CONFIGURATION: linéaire
(ii) TYPE DE MOLECULE: ADNc
(iii) HYPOTHETIQUE: NON
(iii) ANTI-SENS: NON
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°4

25

CCATGGTAGA GTAGGTATCT ATTCGGTAT GAAACTCCA ACTCTGTAG AGGTCTCGAG 60
AAAT 64

30 (6) INFORMATION POUR LA SEQ ID NO: 5

- 35 (i) CARACTERISTIQUES DE LA SEQUENCE:
(A) LONGUEUR: 68 paires de bases
(B) TYPE: acide nucléique
(C) NOMBRE DE BRINS: simple
(D) CONFIGURATION: linéaire
(ii) TYPE DE MOLECULE: ADNc
(iii) HYPOTHETIQUE: NON
(iii) ANTI-SENS: NON
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°5

40

CTAGATTCT CGAGACCTCT ACAAGATGTG GAGTTTTCAT ACCGAAATAG ATACCTACTC 60
TACCATGG 68

45

26

- (7) INFORMATION POUR LA SEQ ID NO: 6:
 (i) CARACTERISTIQUES DE LA SEQUENCE:
 (A) LONGUEUR: 34 paires de bases
 (B) TYPE: acide nucléique
5 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION: linéaire
 (ii) TYPE DE MOLECULE: ADNc
 (iii) HYPOTHETIQUE: NON
 (iii) ANTI-SENS: NON
10 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°6

CCATGGTAGA GTAGGTATCT ATTCGGTAT GAAA

34

- 15 (8) INFORMATION POUR LA SEQ ID NO: 7:
 (i) CARACTERISTIQUES DE LA SEQUENCE:
 (A) LONGUEUR: 40 paires de bases
 (B) TYPE: acide nucléique
 (C) NOMBRE DE BRINS: simple
20 (D) CONFIGURATION: linéaire
 (ii) TYPE DE MOLECULE: ADNc
 (iii) HYPOTHETIQUE: NON
 (iii) ANTI-SENS: NON
 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°7

25

CTAGTTTCAT ACCGAAATAG ATACCTACTT CTTACCATGG

40

- 30 (9) INFORMATION POUR LA SEQ ID NO: 8:
 (i) CARACTERISTIQUES DE LA SEQUENCE:
 (A) LONGUEUR: 52 paires de bases
 (B) TYPE: acide nucléique
 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION: linéaire
35 (ii) TYPE DE MOLECULE: ADNc
 (iii) HYPOTHETIQUE: NON
 (iii) ANTI-SENS: NON
 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°8

40 ACGGGATGAA GGAAGGCC ATGGTAGAGT AGGTATCTAT TTCGGTATGA AA

52

- (10) INFORMATION POUR LA SEQ ID NO: 9:
 (i) CARACTERISTIQUES DE LA SEQUENCE:
45 (A) LONGUEUR: 61 paires de bases
 (B) TYPE: acide nucléique
 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION: linéaire
 (ii) TYPE DE MOLECULE: ADNc
50 (iii) HYPOTHETIQUE: NON
 (iii) ANTI-SENS: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°9

TCGATTTCAT TACCGAAATA GATACCTACT CTACCATGGG CCTTCCCTTC ATCCCGTAGC 60
T 61

5

(11) INFORMATION POUR LA SEQ ID NO: 10:

- 10 (i) CARACTERISTIQUES DE LA SEQUENCE:
(A) LONGUEUR: 34 paires de bases
(B) TYPE: acide nucléique
(C) NOMBRE DE BRINS: simple
(D) CONFIGURATION: linéaire
(ii) TYPE DE MOLECULE: ADNc
(iii) HYPOTHETIQUE: NON
15 (iii) ANTI-SENS: NON
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°10

ACATGGTAGA GTAGGTATCT ATTCGGTAT GAAA

34

20 (12) INFORMATION POUR LA SEQ ID NO: 11:

- (i) CARACTERISTIQUES DE LA SEQUENCE:
(A) LONGUEUR: 42 paires de bases
(B) TYPE: acide nucléique
(C) NOMBRE DE BRINS: simple
25 (D) CONFIGURATION: linéaire
(ii) TYPE DE MOLECULE: ADNc
(iii) HYPOTHETIQUE: NON
(iii) ANTI-SENS: NON
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°11

30

TCGATTTCAT ACCGAAATAG ATACCTACTC TACCATGTAG CT

42

(13) INFORMATION POUR LA SEQ ID NO: 12:

- 35 (i) CARACTERISTIQUES DE LA SEQUENCE:
(A) LONGUEUR: 21 paires de bases
(B) TYPE: acide nucléique
(C) NOMBRE DE BRINS: simple
(D) CONFIGURATION: linéaire
(ii) TYPE DE MOLECULE: ADNc
40 (iii) HYPOTHETIQUE: NON
(iii) ANTI-SENS: NON
(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 12:

GATCCATAGA AGGTCGACTA G

21

45

(14) INFORMATION POUR LA SEQ ID NO: 13:

- (i) CARACTERISTIQUES DE LA SEQUENCE:
(A) LONGUEUR: 21 paires de bases
(B) TYPE: acide nucléique

28

- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- (iii) HYPOTHETIQUE: NON
- 5 (iii) ANTI-SENS: NON
- (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°13

CTAGGTATCT TCCAGCTGAT C

21

- 10 (15) INFORMATION POUR LA SEQ ID NO: 14:
 - (i) CARACTERISTIQUES DE LA SEQUENCE:
 - (A) LONGUEUR: 19 paires de bases
 - (B) TYPE: acide nucléique
 - (C) NOMBRE DE BRINS: simple
 - 15 (D) CONFIGURATION: linéaire
 - (ii) TYPE DE MOLECULE: ADNc
 - (iii) HYPOTHETIQUE: NON
 - (iii) ANTI-SENS: NON
 - (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°14

20

CAGAATCGAA GGTAGAGCC

19

- (16) INFORMATION POUR LA SEQ ID NO: 15:
 - (i) CARACTERISTIQUES DE LA SEQUENCE:
 - (A) LONGUEUR: 30 paires de bases
 - (B) TYPE: acide nucléique
 - (C) NOMBRE DE BRINS: simple
 - (D) CONFIGURATION: linéaire
 - (ii) TYPE DE MOLECULE: ADNc
 - 30 (iii) HYPOTHETIQUE: NON
 - (iii) ANTI-SENS: NON
 - (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°15

TCGAGGCTCT ACCTTCGATC GAGGGTAGCT

30

- 35 (17) INFORMATION POUR LA SEQ ID NO: 16:
 - (i) CARACTERISTIQUES DE LA SEQUENCE:
 - (A) LONGUEUR: 24 paires de bases
 - (B) TYPE: acide nucléique
 - 40 (C) NOMBRE DE BRINS: simple
 - (D) CONFIGURATION: linéaire
 - (ii) TYPE DE MOLECULE: ADNc
 - (iii) HYPOTHETIQUE: NON
 - (iii) ANTI-SENS: NON
 - 45 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°16

ACCCTCGATC GAAGGTAGAT CTCCA

24

- 50 (18) INFORMATION POUR LA SEQ ID NO: 17:
 - (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 33 paires de bases
 (B) TYPE: acide nucléique
 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION: linéaire
- 5 (ii) TYPE DE MOLECULE: ADNc
 (iii) HYPOTHETIQUE: NON
 (iii) ANTI-SENS: NON
 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°17
- 10 TCGATGGAGA TCTACCTTCG ATCGAGGGTA GCT 33
- (19) INFORMATION POUR LA SEQ ID NO: 18:
 (i) CARACTERISTIQUES DE LA SEQUENCE:
 (A) LONGUEUR: 31 paires de bases
 15 (B) TYPE: acide nucléique
 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION: linéaire
 (ii) TYPE DE MOLECULE: ADNc
 (iii) HYPOTHETIQUE: NON
 20 (iii) ANTI-SENS: NON
 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°18
- GTCCCGGATG GAGCGCGTAC TTAGAGAGAA T 31
- 25 (20) INFORMATION POUR LA SEQ ID NO: 19:
 (i) CARACTERISTIQUES DE LA SEQUENCE:
 (A) LONGUEUR: 39 paires de bases
 (B) TYPE: acide nucléique
 (C) NOMBRE DE BRINS: simple
 30 (D) CONFIGURATION: linéaire
 (ii) TYPE DE MOLECULE: ADNc
 (iii) HYPOTHETIQUE: NON
 (iii) ANTI-SENS: NON
 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N°19
- 35 TCGACAGGGC CTACCTCGCG CATGAATCTC TCTTAAGCT 39
- (21) INFORMATION POUR LA SEQ ID NO: 20:
 (i) CARACTERISTIQUES DE LA SEQUENCE:
 40 (A) LONGUEUR: 11 acides aminés
 (B) TYPE: acide aminé
 (D) CONFIGURATION: linéaire
 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 20
- 45 His Gly Arg Val Gly Ile Tyr Phe Gly Met Lys

(22) INFORMATION POUR LA SEQ ID NO: 21:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 4 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

5

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 21

Ile Glu Gly Arg

10 (23) INFORMATION POUR LA SEQ ID NO: 22:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 4 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

15

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 22

Asp Glu Gly Lys

20 (24) INFORMATION POUR LA SEQ ID NO: 23:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 8 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

25

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 23

Pro Ser Ile Glu Gly Arg Ser Pro

5

30 (25) INFORMATION POUR LA SEQ ID NO: 24:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 9 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

35

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 24

Arg Met Glu Arg Val Leu Arg Glu Asn

5

REVENDICATIONS

1. Polypeptide recombinant comportant au moins une partie active dérivée d'un polypeptide, naturel ou synthétique, biologiquement actif, génétiquement insérée dans une albumine ou un de ses variants ou dérivés.
- 5 2. Polypeptide selon la revendication 1 caractérisé en ce que le polypeptide biologiquement actif possède une activité thérapeutique et est d'origine humaine.
3. Polypeptide selon la revendication 2 caractérisé en ce que le polypeptide ayant une activité thérapeutique est choisi parmi tout ou partie des enzymes, des inhibiteurs d'enzymes, des antigènes, des anticorps, des hormones, des récepteurs,
10 des facteurs de la coagulation, des interférons, des cytokines, des facteurs de croissance et/ou de différenciation, des facteurs impliqués dans la génèse/résorption des tissus osseux, des facteurs chimiotactiques, des facteurs de motilité ou de migration cellulaire, des facteurs cytostatiques, des facteurs bactéricides ou antifongiques, ou des molécules adhésives plasmatiques, interstitielles ou des
15 matrices extracellulaires.
4. Polypeptide selon l'une des revendications 1 à 3 caractérisé en ce que le polypeptide ayant une activité thérapeutique est choisi parmi toute séquence peptidique antagoniste ou agoniste d'interactions moléculaires et/ou cellulaires impliquées dans les pathologies des compartiments circulatoires et interstitiels.
- 20 5. Polypeptide selon l'une des revendications 1 à 4 caractérisé en ce que la partie active présente une structure choisie parmi :
 - (a) la structure peptidique entière ou,
 - (b) un fragment de (a) ou une structure dérivée de (a) par modification structurale (mutation, substitution, addition et/ou délétion d'un ou plusieurs résidus) et
25 conservant une activité thérapeutique.
6. Polypeptide selon l'une des revendications 1 à 5 caractérisé en ce que la partie active est insérée strictement à l'intérieur de l'albumine ou entourée de séquence de jonctions.

7. Polypeptide selon l'une des revendications 1 à 6 caractérisé en ce que la partie active est insérée de préférence au niveau des régions de l'albumine présumées former des régions exposées à la surface de la molécule.
8. Polypeptide selon l'une des revendications 1 à 7 caractérisé en ce que la
5 partie active est insérée au niveau de la région 5 s'étendant du résidu 57 à 62 de l'albumine.
9. Polypeptide selon l'une des revendications 1 à 7 caractérisé en ce que la partie active est insérée au niveau de la région 8 s'étendant du résidu 103 à 120 de l'albumine.
10. Polypeptide selon l'une des revendications 1 à 7 caractérisé en ce que la
10 partie active est insérée au niveau de la région 13 s'étendant du résidu 178 à 200 de l'albumine.
11. Polypeptide selon l'une des revendications 1 à 7 caractérisé en ce que la
15 partie active est insérée au niveau de la région du résidu 415 au résidu 425 délimitée par les hélices h2 et h3 du domaine III dans l'albumine.
12. Polypeptide selon l'une des revendications 1 à 12 caractérisé en ce que la partie active y est insérée sous forme unique ou multiple.
13. Polypeptide selon la revendication 12 caractérisé en ce que la partie active
20 est répétée plusieurs fois au même endroit et/ou dans des régions différentes de l'albumine.
14. Polypeptide selon l'une des revendications 1 à 13 caractérisé en ce que les parties actives insérées sont de nature différente.
15. Variant d'une séquence nucléotidique codant pour l'albumine ou un de ses variants ou dérivés intégrant au moins un site de restriction unique non naturel.
16. Séquence nucléotidique codant pour un polypeptide selon l'une quelconque
25 des revendications 1 à 14.
17. Séquence nucléotidique selon la revendication 16 caractérisée en ce qu'elle comprend une séquence "leader" permettant la sécrétion du polypeptide exprimé.

18. Cassette d'expression comprenant une séquence nucléotidique selon l'une des revendications 16 ou 17 sous le contrôle d'une région d'initiation de la transcription et éventuellement d'une région de terminaison de la transcription.
19. Plasmide autorépliquatif comportant une cassette d'expression selon la
5 revendication 18.
20. Cellule recombinante eucaryote ou procaryote dans laquelle a été inséré une séquence nucléotidique selon l'une des revendications 16 ou 17 ou une cassette d'expression selon la revendication 18 ou un plasmide selon la revendication 19.
21. Cellule recombinante selon la revendication 20 caractérisée en ce qu'il s'agit
10 d'une levure, d'une cellule animale, d'un champignon ou d'une bactérie.
22. Cellule recombinante selon la revendication 21 caractérisée en ce qu'il s'agit d'une levure.
23. Cellule recombinante selon la revendication 22 caractérisée en ce qu'il s'agit d'une levure du genre Saccharomyces ou Kluyveromyces.
- 15 24. Procédé de préparation d'un polypeptide tel que défini dans l'une des revendications 1 à 14 caractérisé en ce que l'on cultive une cellule recombinante selon l'une des revendications 20 à 23 dans des conditions d'expression, et on récupère le polypeptide produit.
25. Composition pharmaceutique comprenant un ou plusieurs polypeptides selon
20 l'une quelconque des revendications 1 à 14.
26. Composition pharmaceutique comprenant une séquence nucléotidique selon l'une quelconque des revendications 16 à 17 utilisable en thérapie génique.

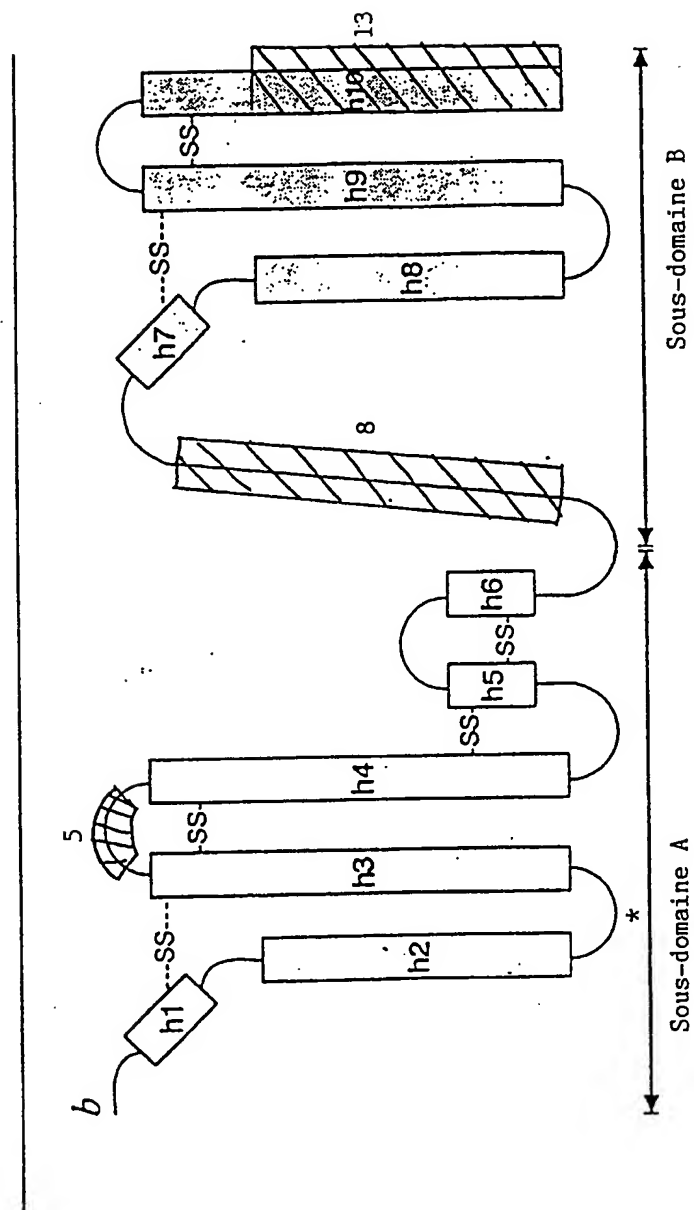


Figure 1

2/6

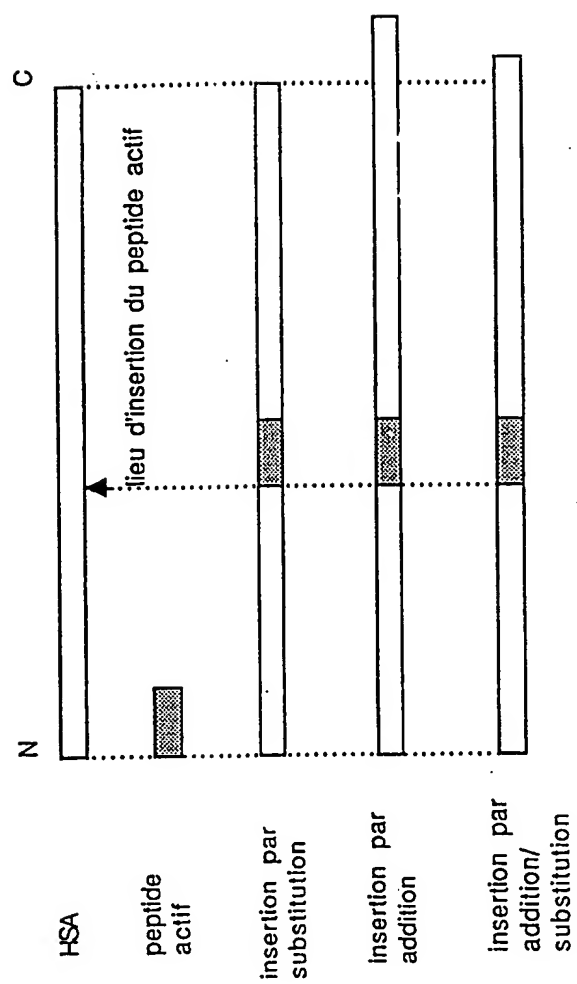


Figure 2

3/6

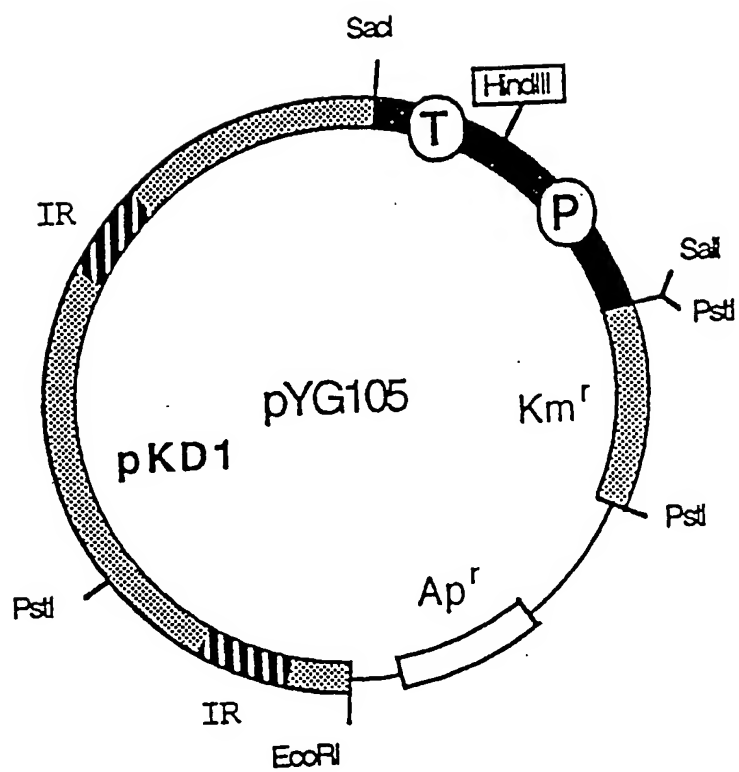


Figure 3

5/6

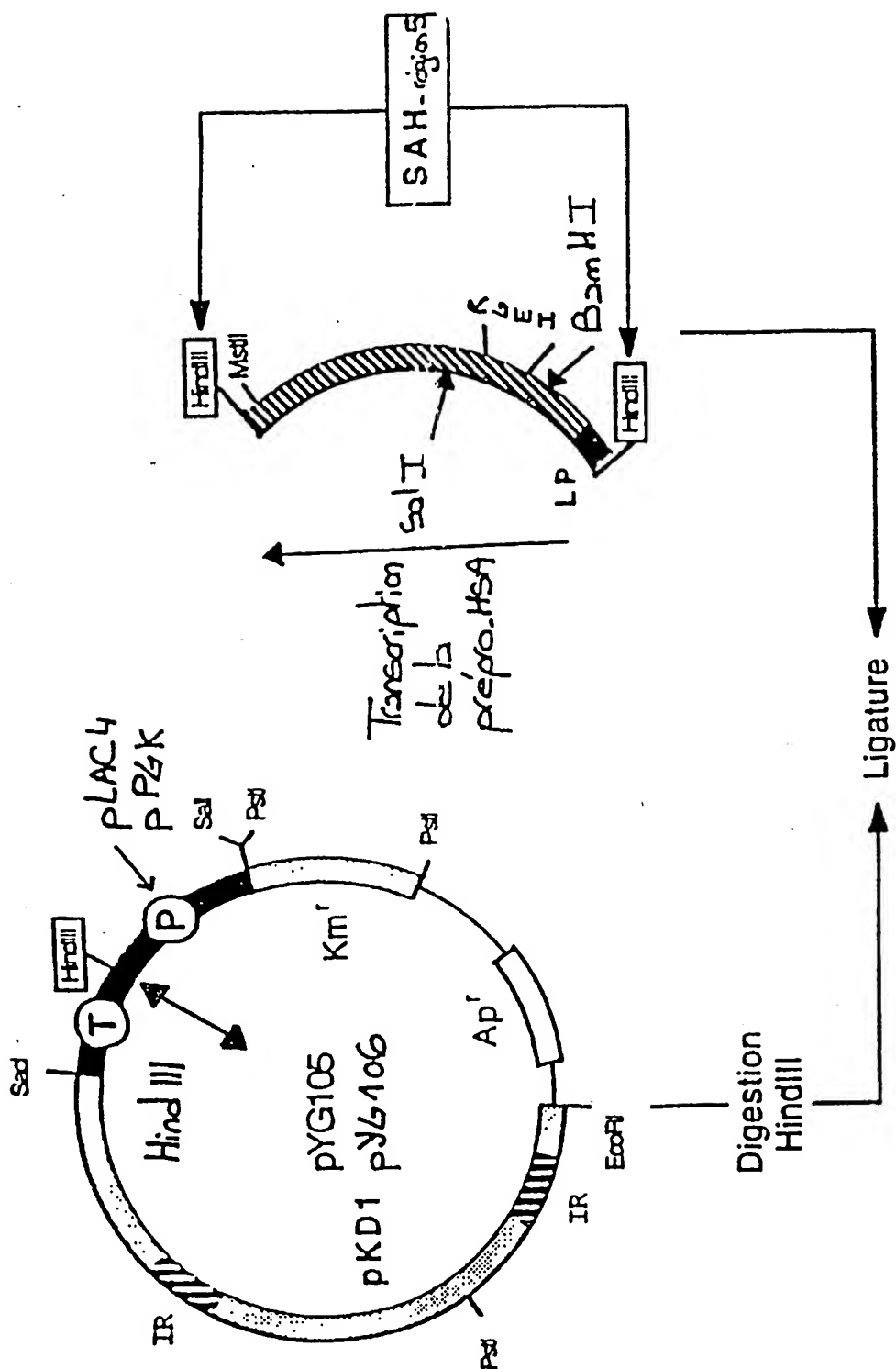


Figure 5

6/6

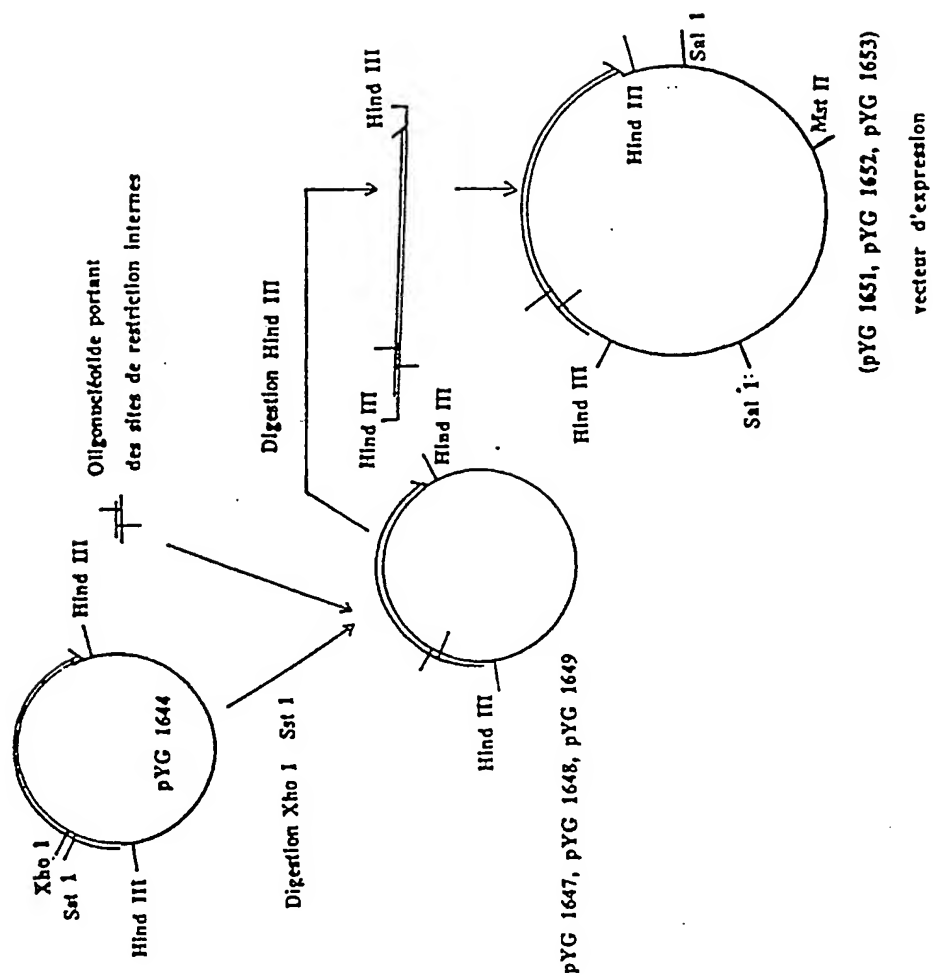


Figure 6

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/FR 95/00520

A. CLASSIFICATION OF SUBJECT MATTER IPC 6 C12N15/62 C12N15/14 C12N15/12 C12N15/60 C12N15/80 C12N15/81 C12N1/19 C12N1/15 A61K38/38 A61K38/17 C07K14/765 //C12N9/88, C12N1/21, C12N5/10,		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 6 C07K C12N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO-A-93 15199 (RHONE-POULENC-RORER S.A) 5 August 1993 see the whole document ---	1-26
A	BIOTECHNOLOGY, vol. 7, September 1989 pages 929-932, J. VANDEKERCKHOVE ET AL. 'Enkephalins produced in transgenic plants using modified 2S seed storage proteins' see the whole document ---	1-3, 5, 6, 12, 16-20, 25
A	EP-A-0 399 666 (DELTA BIOTECHNOLOGY LIMITED) 28 November 1990 * see the whole document, in particular examples 2 and 3 * --- -/-	1-5, 7, 10, 12, 15-26
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : 'A' document defining the general state of the art which is not considered to be of particular relevance 'E' earlier document but published on or after the international filing date 'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) 'O' document referring to an oral disclosure, use, exhibition or other means 'P' document published prior to the international filing date but later than the priority date claimed 'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention 'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone 'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. 'Z' document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
14 August 1995		23.08.95
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016		Authorized officer Gac, G

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/FR 95/00520

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO-A-93 15200 (RHONE-POULENC-RORER S.A.) 5 August 1993 see the whole document ---	1-26
A	WO-A-93 15211 (RHONE-POULENC RORER S.A.) 5 August 1993 see the whole document ---	1-26
A	J. MOL. BIOL., vol. 184, no. 4, 20 August 1985 pages 547-564, JUBIER-MAURIN ET AL. 'Comparative study of the L1 family in the genus Mus : possible role of retroposition and conversion events in its concerted evolution' * pages 547, 552 * * page 555 right column * -----	1,5,6, 12,16

INTERNATIONAL SEARCH REPORT

International Application No
PCT/FR 95/00520

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9315199	05-08-93	FR-A- 2686899	06-08-93
		EP-A- 0624195	17-11-94
		FI-A- 943563	29-07-94
		JP-T- 7503368	13-04-95

EP-A-0399666	28-11-90	DE-T- 69002395	25-11-93
		EP-A- 0407008	09-01-91
		ES-T- 2060033	16-11-94
		WO-A- 9013306	15-11-90

WO-A-9315200	05-08-93	FR-A- 2686901	06-08-93
		CA-A- 2126092	05-08-93
		EP-A- 0625199	23-11-94
		FI-A- 943565	29-07-94
		JP-T- 7503369	13-04-95
		NO-A- 942840	29-09-94

WO-A-9315211	05-08-93	FR-A- 2686900	06-08-93
		CA-A- 2125979	05-08-93
		EP-A- 0624200	17-11-94
		FI-A- 943564	29-07-94
		JP-T- 7503844	27-04-95
		NO-A- 942858	01-08-94

RAPPORT DE RECHERCHE INTERNATIONALE

 Dem: Internationale No
 PCT/FR 95/00520

A. CLASSEMENT DE L'OBJET DE LA DEMANDE

CIB 6	C12N15/62	C12N15/14	C12N15/12	C12N15/60	C12N15/80
	C12N15/81	C12N1/19	C12N1/15	A61K38/38	A61K38/17
	C07K14/765		//C12N9/88, C12N1/21, C12N5/10,		

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 6 C07K C12N

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés)

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	WO-A-93 15199 (RHONE-POULENC-RORER S.A) 5 Août 1993 voir le document en entier ---	1-26
A	BIOTECHNOLOGY, vol. 7, Septembre 1989 pages 929-932, J. VANDEKERCKHOVE ET AL. 'Enkephalins produced in transgenic plants using modified 2S seed storage proteins' voir le document en entier ---	1-3, 5, 6, 12, 16-20, 25
A	EP-A-0 399 666 (DELTA BIOTECHNOLOGY LIMITED) 28 Novembre 1990 *le document en entier, surtout exemples 2 et 3 * --- -/-	1-5, 7, 10, 12, 15-26

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

* Catégories spéciales de documents cités:

- *A* document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- *E* document antérieur, mais publié à la date de dépôt international ou après cette date
- *L* document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- *O* document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- *P* document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

T document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

X document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

Y document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

Z document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

14 Août 1995

Date d'expédition du présent rapport de recherche internationale

23.08.95

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+ 31-70) 340-3016

Fonctionnaire autorisé

Gac, G

RAPPORT DE RECHERCHE INTERNATIONALE

Dem Internationale No
PCT/FR 95/00520

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	WO-A-93 15200 (RHONE-POULENC-RORER S.A.) 5 Août 1993 voir le document en entier ---	1-26
A	WO-A-93 15211 (RHONE-POULENC RORER S.A.) 5 Août 1993 voir le document en entier ---	1-26
A	J. MOL. BIOL., vol. 184, no. 4, 20 Août 1985 pages 547-564, JUBIER-MAURIN ET AL. 'Comparative study of the L1 family in the genus Mus : possible role of retroposition and conversion events in its concerted evolution' * pages 547, 552 * * page 555 colonne de droite * -----	1,5,6, 12,16

RAPPORT DE RECHERCHE INTERNATIONALE

Dem Internationale No
PCT/FR 95/00520

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO-A-9315199	05-08-93	FR-A- 2686899	06-08-93
		EP-A- 0624195	17-11-94
		FI-A- 943563	29-07-94
		JP-T- 7503368	13-04-95
EP-A-0399666	28-11-90	DE-T- 69002395	25-11-93
		EP-A- 0407008	09-01-91
		ES-T- 2060033	16-11-94
		WO-A- 9013306	15-11-90
WO-A-9315200	05-08-93	FR-A- 2686901	06-08-93
		CA-A- 2126092	05-08-93
		EP-A- 0625199	23-11-94
		FI-A- 943565	29-07-94
		JP-T- 7503369	13-04-95
		NO-A- 942840	29-09-94
WO-A-9315211	05-08-93	FR-A- 2686900	06-08-93
		CA-A- 2125979	05-08-93
		EP-A- 0624200	17-11-94
		FI-A- 943564	29-07-94
		JP-T- 7503844	27-04-95
		NO-A- 942858	01-08-94

COMMONWEALTH LANGUAGE SERVICES, LTD.*Alison Sondhaus Carroll, President**6044 Little Falls Rd., Arlington, Virginia 22207**tel: (703) 534-7649 fax: (703) 534-4191**email: alisonscarroll@juno.com***MEMORANDUM**

To: Lara Kelley, esq./Brian Fish

From: Alison Sondhaus Carroll

Date: October 9, 2001

Re: certification

International patent application No. PCT/FR95/00520

International filing date: April 20, 1995

Applicant: Rhone-Poulenc Rorer, S.A.

Title: "Biologically Active Polypeptides Inserted into an Albumin"

Ref no.: 06832.0004-02000

Translated from the French by Commonwealth Language Services, Ltd.

PCT

World Intellectual Property Organization
International Office

International Patent Application published in accordance with the Patent Convention Treaty (PCT)

(51) International patent classification ⁶:
C12N 15/62, 15/14, 15/12, 15/50, 15/80,
15/81, 1/19, 1/15, A61K 38/38, 38/17, A1
C07K 14/765 // C12N 9/88, 1/21, 5/10

(11) International Publication No.: WO 95/30759
(43) International Publication Date: 11/16/95

(21) International application no: PCT/FR95/00520

(81) Designated countries: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, IE, IT, LU, MC, NL, PT, SE)¹

(22) International filing date: 4/20/95

Published

(30) Information related to priority:
94/05616 5/6/94 FR

With international search report

(71) Applicant (*for all designated countries except US*):
RHONE-POULENC RORER S.A. (FR/FR); 20,
Avenue Raymond-Aron, F-92160 Antony (FR)

(72) Inventors; and
(75) Inventors/Applicants (*US only*): BECQUART,
Jerome (FR/FR), 8 rue Maublanc, F-75015 Paris (FR);
CONSEILLER, Emmanuel (FR/FR), 10 rue de Plelo,
F-75015 Paris (FR); GUITTON, Jean-Dominique (FR/FR),
74, rue Dunois, F-75013 Paris (FR); HARDY, Florence (FR/FR),
6, square Villaret-de-Joyeuse, F-75017 Paris (FR); YEH,
Patrice (FR/FR), 11 bis rue Lacepede, F-75005 Paris (FR)

(74) Agent: LE COUPANEC, Pascale, Rhone-Poulenc Rorer
S.A., Patent Department, 20, avenue Raymond-Aron, F-92160
Antony (FR)

(54) Title: BIOLOGICALLY ACTIVE POLYPEPTIDES INSERTED INTO AN ALBUMIN

(54) [title in French]

(57) Abstract

Biologically active recombinant polypeptides essentially consisting of at least one active portion derived from a natural or artificial biologically active peptide and inserted into an albumin or albumin variant, the preparation thereof, and pharmaceutical compositions containing same, are disclosed.

(57) [abstract in French]

¹ AT = Austria, BE = Belgium, CH = Switzerland, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = United Kingdom, IE = Ireland, IT = Italy, LU = Luxemburg, MC = Monaco, NL = Netherlands, PT = Portugal, SE = Sweden

BIOLOGICALLY ACTIVE POLYPEPTIDES INSERTED INTO AN ALBUMIN

The present invention relates to new biologically active polypeptides, their preparation and pharmaceutical compositions containing them.

Although possessing one or several potential therapeutic activities, numerous polypeptides unfortunately cannot be exploited pharmaceutically. This may be for different reasons, in particular their weak stability *in vivo*, their complex or fragile structure, the difficulty of producing them on an acceptable industrial scale, etc. Likewise, certain polypeptides do not give the expected results *in vivo*, due to problems of administration, conditioning, pharmacokinetics, etc.

The aim of the present invention is precisely to overcome these drawbacks.

It is particularly aimed at the development of artificial, biologically active proteins permitting optimal exploitation of the biological properties of these polypeptides for therapeutic use.

The applicant accordingly provides evidence that using genetic methods, it is possible to insert any active structure derived from a biologically active protein into another protein structure consisting of albumin, without altering the said biological properties. In an unexpected way, human serum albumin allows the active structure to be efficiently introduced at its sites of interaction, and increases the plasmatic stability of the recombinant polypeptide embodied in the invention.

More precisely, the present invention relates to a recombinant polypeptide containing at least one active portion derived from a natural or artificial biologically active polypeptide inserted genetically into an albumin or one of its variants or derivatives.

According to the present invention, a variant of albumin is understood to designate any protein with long plasma half-life obtained by modification of a gene coding for a given isomorph of the human serum albumin by genetic

engineering techniques. It is also understood to designate any macromolecule with long plasma half-life obtained by *in vitro* modification of the protein coded for by such genes. (Modification is understood to mean any mutation, substitution, deletion, addition or modification of a genetic and/or chemical nature.) Albumin being very polymorphic, numerous natural variants have been identified and listed [Weitkamp L.R. et al., Ann. Hum. Genet. 37 (1973) 219].

Derivatives of albumin most particularly refers to molecules consisting wholly or partly of albumin, linked with at least one polypeptide sequence resulting from a natural or artificial gene, whether or not it has biological activity itself.

In the description that follows, the different types of albumins clarified above are commonly designated by the term "albumin."

In the sense of the present invention, active portion is understood to mean a portion having an activity which can be either direct (treatment of diseases, diagnosis, biological research, markers) or indirect (for example, usable in preventing diseases, in the design of vaccines, in medical imaging techniques, etc).

The active portions of biologically active polypeptides inserted according to the invention are preferably of therapeutic interest.

Polypeptides having therapeutic activity may or may not be of human origin.

As representative of polypeptides of non-human origin, one may cite peptides or their derivatives having properties potentially useful in the pathologies of blood or interstitial compartments, such as hirudine, trigramine, antistatine, the anticoagulant peptides of ticks (TAP), arietine, applagine, etc.

According to a preferred embodiment of the invention, the polypeptide having a therapeutic activity is a polypeptide of human origin or a molecular variant. For example, it may consist, wholly or partly, of an enzyme, an enzyme inhibitor, an antigen, an antibody, a hormone, a receptor, a factor

3.

involved in the control of coagulation, an interferon, a cytokine [the interleukins, but also the natural variant antagonists of their attachments to receptors, the SIS (small induced secreted)-type cytokines and, for example, the inflammatory proteins of macrophages (MIPs), etc], a growth and/or differentiation factor [and for example the transforming growth factors (TGFs), the differentiation factors of blood cells (erythropoietin, M-CSF, G-CSF, GM-CSF, etc.), insulin and the growth factors that resemble it (IGFs), or also the cellular permeability factors (VPF/VEGF), etc.], a factor implicated in the genesis or resorption of bone tissue (OIF and osteospondine, for example), a factor implicated in cellular motility or migration [for example, the autocrine mobility factor (AMF), the migration stimulating factor (MSF), or also the dispersion factor (scatter factor/hepatocyte growth factor)], a bactericidal or antifungal factor, a chemotactic factor [for example, the platelet factor 4 (PF4), or the monocyte chemoattractant peptides (MCP/MCAF) or neutrophils (NCAF), etc.], a cytostatic factor (for example, the proteins that attach to galactosides), an adhesive molecule, either plasmatic (for example, the von Willebrand factor, fibrinogen, etc.), or interstitial (laminine, tenascene, vitronectine, etc.), or extracellular matrices, or also any peptide sequence antagonistic or agonistic to molecular and/or intercellular interactions implicated in the pathologies of circulatory and interstitial compartments, for example the formation of arterial or venous thrombi, cancerous metastases, tumoral angiogenesis, inflammatory shock, autoimmune diseases, bone and osteo-articular pathologies, etc.

Of course, the active portion of the polypeptides of the invention may consist of the entire biologically active peptide or a structure derived from it, or may also correspond to a non-natural peptide sequence, isolated for example from random peptide banks. (For purposes of simplification, these different possibilities will be covered in what follows by the collective designation "active portion of a biologically active peptide.") In the sense of the present invention, derived structure is understood to mean any polypeptide obtained by modification which retains its biological activity. Modification is understood to mean any mutation, substitution, deletion, addition or modification of genetic and/or chemical nature. Such derivatives may be generated for different purposes, especially

4.

such as increasing the affinity of the molecule for its attachment sites, improving its level of production, increasing its resistance to proteases, enhancing its therapeutic efficacy or reducing its secondary effects, or giving it new biological properties. As an example, the chimeric polypeptides of the invention have pharmacokinetic properties and biological activity usable in the prevention or treatment of diseases.

The polypeptides according to the invention that are particularly advantageous are those in which the active portion has:

- (a) the entire peptide structure, or
- (b) a fragment of (a) or a structure derived from (a) by structural modification (mutation, substitution, addition and/or deletion of one or several residues) and possessing therapeutic activity.

Among the structures of type (b) may be cited in particular the molecules in which certain sites of N- or O-glycosylation have been modified or deleted, molecules in which one or several residues have been substituted, or molecules in which all the cysteine residues have been substituted. One may also cite molecules obtained from (a) by deleting the regions having little or no involvement in the interaction with the linkage sites under consideration, or expressing an undesirable activity, and molecules including supplementary residues with respect to (a), for example, a terminal methionine-N, and/or a secretion signal, and/or a junction peptide.

The subject matter of the present invention is particularly advantageous in regard to active peptides that are too small to form a structural domain, and/or do not possess good stability *in vivo* and/or good bioavailability. The proposed method of insertion according to the invention allows them to be associated with one or more preexisting domains of albumin and thus to benefit from the bioavailability and stability *in vivo* of the latter.

In general, the sizes of the active portions inserted into the albumin vary between three and twenty five amino acid residues. However, sequences of 1 to 100 residues can also be used.

5.

The insertion of an active portion of a peptide into the peptide sequence of albumin is accomplished according to the invention in a way satisfying the following two conditions:

Sufficient accessibility must be preserved to keep intact the biological activity of the active portion inserted into the albumin. Moreover, the structure of the albumin of course cannot be subjected to a destabilization significant enough to be detrimental to the recombinant polypeptide, or chimera.

The sites of insertion are preferably selected within the albumin while adhering to the above requirements.

According to the crystalline structure published by He and Carter (Nature 1992, 358, 209-215), albumin is formed by the repetition of 3 domains, each comprising two subdomains, and it consists of over 67% alpha helices. Each domain is superimposable on the others and is composed of 10 helices designated h1 to h10. The subdomain A consists of helices h1 to h6 and the subdomain B, helices h7 to h10. Each subdomain is formed in a common pattern: h1, h2, h3, h4 for domain A, and h7, h8, h9 and h10 for domain B. The small supplementary helices h5 and h6 are linked by a disulfide bond to subdomain A. Figure 1 schematizes the structure of human serum albumin.

The sites of insertion are preferably located in the regions of the albumin presumably forming the exposed regions on the surface of the molecule; these regions preferentially being loops.

As insertion sites that are particularly suitable for the invention, three regions of the first domain may be mentioned:

- region 5, which extends from residue 57 to 62 and corresponds to a loop joining the helices h3 and h4;
- region 8, comprising the residues 103 to 120 and corresponding to the inter-subdomain zone;
- region 13, between residues 178 and 200, which corresponds to a helix.

According to the invention, helices h2 and h3 of domain III delimit another conceivable insertion region, from residues 419 to 430.

An active portion of the biologically active peptide may be inserted in the albumin peptide sequence according to three different embodiments of the invention:

6.

- A strict insertion consisting of a simple addition of the peptide sequence of interest into the original sequence of the albumin, which is totally conserved.
- The insertion may correspond to a substitution of a portion of the albumin peptide sequence by the peptide sequence corresponding to the active portion of the peptide of interest.
- Finally, an insertion combining an addition of a part of the active peptide sequence and a substitution of a portion of the albumin peptide sequence by the rest of the active portion of the active peptide.

Figure 2 schematically depicts these different modes of insertion.

Of course, the active portion of a biologically active peptide can be repeated several times in the chimera at the same place and/or in different regions of the albumin. Likewise, it is equally possible as another embodiment of the invention to insert different active portions, whether stemming from the same peptide or from different peptides.

Finally, in another embodiment of the invention, an active portion of a peptide may be inserted either within the albumin or by surrounding the sequences of the junction.

In regard to the junction sequences, peptide sequences rich in glycine residues and/or serine residues and/or threonine residues, and/or any amino acid residue described as frequently encountered in the zones of flexibility in proteins, may be especially involved.

The recombinant polypeptides according to the invention prove to be particularly advantageous.

They permit a given biological activity to be maintained in an organism for an extended period. It thus proves to be possible to reduce the administered doses, and in certain cases to potentiate the therapeutic effect, for example by reducing the consecutive secondary effects of a larger dose. Advantageously, they allow the production and utilization of structures derived from biologically active polypeptides that are very small and therefore very specific to a desired effect. Moreover, these recombinant polypeptides are distributed in the organism especially advantageously,

modifying their pharmacokinetic properties and favoring the development of their biological activity and their utilization. They also have the advantage of

7.

being weakly or nonimmunogenic for the organism in which they are used. Finally, the polypeptides of the invention may be expressed (or preferably secreted) by recombinant organisms at levels permitting their industrial use.

Another aim of the invention relates to a process for preparing the chimeric molecules described above. More precisely, this process consists of causing a nucleotide sequence coding for an active portion of the desired polypeptide to be expressed by a eukaryote or prokaryote cellular host, then harvesting the product polypeptide.

Among the eukaryote hosts usable within the scope of the present invention one may cite animal cells, yeasts or fungi. In particular, for yeasts one may cite those of the type Saccharomyces, Kluyveromyces, Pichia, Schwanniomyces or Hansenula. For animal cells, COS, CHO, C127, etc. may be cited. Among the fungi able to be used according to the present invention one may cite in particular Aspergillus ssp, or Trichoderma ssp. As prokaryote hosts, bacteria such as Escherichia coli or those belonging to the type Corynebacterium, Bacillus or Streptomyces would be preferable.

The nucleotide sequences usable within the scope of the present invention can be prepared in different ways. Generally they are obtained by assembling in the reading frame the sequences coding for each of the functional parts of the polypeptide. These can be isolated by techniques of a person skilled in the art, for example, directly from cellular ARN messengers (ARNm), or by recloning from a bank of complementary ADN (ADNc), or again, totally synthetic nucleotide sequences may be involved. Moreover, it is understood that nucleotide sequences may also be subsequently modified, for example, by genetic engineering techniques, to obtain derivatives or variants of the said sequences.

The insertion of this nucleotide sequence coding for the active portion of the polypeptide may or may not be directly achieved, depending on the region chosen for the insertion site in the gene coding for albumin.

The region selected may not actually include the restriction site adequate for the said insertion. According to this hypothesis, it may prove to be of interest to introduce one or several single restriction sites prior to the insertion.

8.

The creation of restriction sites at the level of the region chosen for the insertion site is preferably done by directed mutagenesis according to traditional techniques. However, one can also insert the nucleotide sequence corresponding to "the active portion of the biologically active peptide" directly by directed mutagenesis so that the insertion does not cause particular restriction sites to appear.

In the particular case of region 5 in the gene coding for albumin, the presence of restriction site PvuII permits the direct insertion of the nucleotide sequence. It thus shows itself to be particularly useful as a cloning site for an active portion of a polypeptide which one hopes to insert into the translational [reading] frame within the albumin sequence at the 57th residue level.

In this mode of insertion, which makes use of a single restriction site already existing in the original albumin sequence, the ligation of the sequence coding for the active peptide with the restriction fragment corresponding to the totality of the gene coding for albumin, generates a nucleotide sequence comprising a hybrid gene coding for a strict insertion HAS-type chimeric protein.

In the particular case of zones 419 to 430, the presence of 2 insertion sites (HincII and ArvII) allows the peptides of biological interest to be inserted and/or substituted in the albumin sequence.

With particular reference to regions 8 and 13, the insertion of the said sequence is favored if one or more restriction sites that can be manipulated are created for it beforehand. To be sure, the restriction sites to be created on the sequence level are chosen by taking into account the nature of the preexisting restriction sites. They must lead to a selective insertion.

In the second embodiment, the use of two single restriction sites allows the construction of genes coding for chimeras having the active peptide inserted and/or substituted. The peptide insertion is made by replacement of a narrow fragment of two single restriction sites in the complementary ADN of the albumin, inserted by directed mutagenesis. These two single restriction sites can be Mst I and Kpn I in region 8 and Sst I and Xho I in region 13. The creation of these restriction sites may or may not modify the polypeptide

sequence of human albumin. The subsequent cloning of active peptide in coding phase in this albumin gene may be exclusively the peptide

9.

coding sequence, or may correspond to a blend of the coding sequence of the peptide and the coding sequence of the deleted albumin fragment.

The present invention also aims to protect the variants of the nucleotide sequences coding for the corresponding albumin; that is, integrating at least one single non natural restriction site, not present in the original sequence.

Creation of these restriction sites can additionally serve to advantage in inserting one or several active portions of biologically active polypeptides into the mature protein.

More preferably, in the process that is the subject matter of the invention, the nucleotide sequence is part of an expression cassette comprising a transcription initiation region (promoter region), permitting expression of the nucleotide sequence in the host cells placed under its control and coding for the polypeptides in the invention. This region can arise from promoter regions of genes strongly expressed in the cellular host utilized, the expression being constitutive or regulated. With respect to yeasts, it may be the promoter of the gene for phosphoglycerate kinase (PKG); for glyceraldehyde-3-phosphate dehydrogenase (GPD); lactase (LAC4); the enolases (ENO); the alcohol dehydrogenases (AOH), etc. As to bacteria, it may be the promoter of the right or left genes of the lambda bacteriophage (P_L , P_R), or also the promoters of genes of the operons tryptophane (P_{trp}) or lactose (P_{lac}). Moreover, this control region can be modified, for example by mutagenesis *in vitro*, by introduction of additional control elements or synthetic sequences, or by deletions or substitutions of the original control elements. The expression cassette can equally comprise a region of functional transcription termination in the host considered, positioned immediately ahead of the nucleotide sequence coding for a polypeptide embodied in the invention.

In a preferred embodiment, the polypeptides of the invention result from the expression of a nucleotide sequence in a eukaryote or prokaryote host and the secretion of the expression product of the said sequence into the culture medium. In fact, it is particularly advantageous to be able to obtain the molecules by recombinant pathway directly in the culture medium. In this case, the nucleotide sequence coding for a polypeptide embodied in

10.

the invention is preceded by a "leader" sequence (or signal sequence) directing the polypeptide arising in the secretion pathways of the host utilized. This "leader" sequence may be the natural signal sequence of the biologically active peptide in the case where this is a naturally secreted protein, or may be that of the stabilizing structure, but it may equally well consist of any other functional "leader" sequence, or of an artificial "leader" sequence. The choice of one or another of these sequences is guided in particular by the host utilized. Examples of functional signal sequences include those of the sexual pheromone genes or the "killer" toxins of yeasts.

Besides the expression cassette, one or more markers permitting the selection of the recombinant host can be added, such as for example the gene URA3 of the yeast S.Cerevisiae, or the genes conferring resistance to antibiotics such as geneticine (G418), or any other toxic compound such as certain metallic ions.

The ensemble comprising the expression cassette and the selection marker may be introduced directly into the host cells in question, that is, inserted beforehand in a functionally self-replicating vector. In the first case, sequences homologous to the regions present in the genome of the host cells are preferably added to this ensemble; the said sequences being then positioned on each side of the expression cassette and of the selection gene as a way to increase the integration frequency of the ensemble in the host genome by targeting the integration of the sequences by homologous recombination. In the case where the expression cassette is inserted into a replicative system, a replication system preferred for yeasts of the type Kluyveromyces is derived from the plasmid pKD1, initially isolated from K. drosophilianum. A replication system preferred for the yeasts of type Saccharomyces is derived from plasmid 2 μ of S.Cerevisiae. Furthermore, this expression plasmid may contain all or part of the said replication systems, or may combine elements derived from plasmid pKD1 as well as from plasmid 2 μ .

In addition, the expression plasmids may be shuttle vectors between a bacterial host such as Escherichia coli and the chosen cellular host. In this case, a replication origin and a selection marker functioning in the bacterial host are required. It is also possible to position the restriction sites of

11.

around the bacterial sequences and only on the expression vector; this makes it possible to delete [these sequences] by cutting and religation *in vitro* of the truncated vector before transformation of the host cells, which may result in an increased number of copies and an enhanced stability of the expression plasmids in the said hosts. For example, these restriction sites may correspond to sequences such as 5'-GGCCNNNNNGGCC-3' (SfiI) (SEQ ID N°1) or 5'-GCGGCCGC-3' (NotI) (SEQ ID N°2), inasmuch as these sites are extremely rare and generally absent from an expression vector.

After construction of such vectors or expression cassettes, they are introduced into the host cells retained according to the traditional techniques described in the literature. In this regard, any method permitting the introduction of a foreign ADN into a cell may be used. This includes in particular transformation, electroporation, conjugation or any other technique known to a person skilled in the art. As an example for the hosts of yeast type, the different strains of Kluyveromyces utilized have been transformed by treating the intact cells in the presence of lithium acetate and polyethylene glycol, according to the technique described by Ito et al. [J. Bacteriol. 153 (1983) 183]. The transformation technique described by Durrens et al. [Curr. Genet. 18 (1990) 7] utilizing ethylene glycol and dimethylsulfoxide has also been used. It is also possible to transform yeasts by electroporation, according to the method described by Karube et al. [FEBS Letters 182 (1985) 90]. An alternative protocol is also described in detail in the examples that follow.

After the selection of transformed cells, the cells expressing the said polypeptides are inoculated and the said polypeptides may be recovered, either during cellular growth for the "ongoing" processes, or at the end of growth for batch cultures. The polypeptides that are the subject matter of the present invention are purified using the supernatant of the culture, with the aim of determining their molecular, pharmacokinetic and biological nature.

A preferred system of expression of the polypeptides according to the invention consists of using the yeasts of the Kluyveromyces type as cellular hosts, transformed by certain vectors derived from the extrachromosomal replicon pKD1 initially isolated from K.marxianus var. drosophilianum.

These yeasts, in particular K.lactis and K.Fragilis, are generally capable of replicating the said vectors in a stable manner and furthermore have the

12.

advantage of being included in the G.R.A.S. (Generally Recognized As Safe) list of organisms. Preferred yeasts are preferably the industrial strains of the Kluyveromyces type capable of replicating in stable fashion the said plasmids derived from plasmid pKD1 and in which a selection marker and expression cassette have been inserted permitting the secretion of a high level of the polypeptides according to the invention.

The present invention also relates to the nucleotide sequences coding for the chimeric polypeptides described earlier, as well as the recombinant cells, eukaryotes or prokaryotes, comprising such sequences.

The present invention also relates to the application of the polypeptides according to the present invention as medicines. More particularly, the aim of the invention is any pharmaceutical composition comprising one or several polypeptides or nucleotide sequences such as are described earlier. These pharmaceutical compositions can be presented in various formulations. They may particularly involve nanoparticles having surfaces on which the polypeptides according to the invention are present. This type of formulation is utilized more particularly to achieve directed targeting of the active [molecule]. Of course, the nucleotide sequences can be used in gene therapy.

The present invention will be described more completely with the aid of the examples that follow, which should be considered as illustrative and not limiting.

LIST OF FIGURES

The representations of plasmids indicated in the following figures are not drawn to scale and only the restriction sites that are important for understanding the clonings carried out are indicated.

Figure 1: Schematic representation of the domain I of albumin with location of the insertion sites according to the invention. * signals the location of the h2-h3 insertion site in domain III. The numbers 5, 8 and 13 identify the corresponding insertion zones.

13.

Figure 2: Schematic representation of different modes of insertion of an active peptide in the structure of albumin.

Figure 3: Plasmid pYG105.

Figure 4: Modification of region 5 of albumin following the insertion of the sequence coding for IEGR, as described in example 8.1. The modifications are represented in bold type. The placement of a modified or introduced restriction site is indicated by a horizontal line, and the position of an enzyme cutting with a vertical line.

Figure 5: Cloning strategy for insertion of IEGR in region 5 (example 8.1).

Figure 6: Cloning strategy for insertion of an active peptide in region 13 of albumin.

EXAMPLES

GENERAL CLONING TECHNIQUES

The traditional methods used in molecular biology, such as the preparative extractions of plasma ADN, centrifugation of plasma ADN in a cesium chloride gradient, electrophoresis on agar gels or acrylamide, purification of ADN fragments by electroelution, extraction of proteins in phenol or phenol-chloroform, precipitation of ADN in saline medium by ethanol or isopropanol, transformation in Escherichia coli, etc., are well known to a person skilled in the art and are abundantly described in the literature [Maniatis T. et al., "Molecular Cloning, a Laboratory Manual," Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982; Ausubel F.M. et al. (eds), "Current Protocols in Molecular Biology," John Wiley & Sons, New York, 1987.]

Restriction enzymes were supplied by New England Biolabs (Biolabs), Bethesda Research Laboratories (BRL) or Amersham, and are used according to the recommendations of the suppliers.

Plasmids of type pBR322, pUC and the M13 series phages are commercial in origin (Bethesda Research Laboratories).

For the ligations, the ADN fragments are separated according to size by electrophoresis on agar gel or acrylamide, extracted with phenol or a

14.

phenol/chloroform mixture, precipitated with ethanol and incubated in the presence of the ADN ligase of phage T4 (Biolabs) according to the supplier's recommendations.

The filling of prominent 5' extremities is achieved by the Klenow fragment of E. coli ADN polymerase (Biolabs), used according to the supplier's specifications. The prominent 3' extremities are destroyed in the presence of the ADN polymerase of T4 phage (Biolabs) according to the manufacturer's recommendations. The prominent 5' extremities are destroyed by a treatment brought about with S1 nuclease.

Directed mutagenesis *in vivo* with synthetic oligodeoxynucleotides is achieved according to the method developed by Taylor et al. [Nucleic Acids Res. 13 (1985) 8749-8784], using the kit distributed by Amersham. In the context of the present invention, this technique is utilized in particular to create single restriction sites for use in a subsequent insertion.

Enzymatic amplification of ADN fragments by the technique called PCR [Polymerase-catalyzed Chain Reaction, Saiki R.K. et al., Science 230 (1985) 1350-1354; Mullis K.B. and Faloona F.A., Math. Enzym. 155 (1987) 335-350] is carried out with the use of a "DNA thermal cycler" (Perkin Elmer Cetus) according to the manufacturer's specifications.

The nucleotide sequences are verified using the method developed by Sanger et al. [Proc. Natl. Acad. Sci. USA, 74 (1977) 5463-5487], using the kit distributed by Amersham.

K.lactis is transformed using the ADN from the expression plasmids of the proteins of the present invention by any technique known to a person skilled in the art. An example is given in the text.

Except where otherwise indicated, the bacterial strains used are E.coli MC1060 (lacIPOZYA, X74, galU, galK, strA¹), or E.coli TG1 (lac, proA,B, supE, thi, hsdD5/ F'traD36, proA⁺B⁺, lacI^Q, lacZ, M15).

The yeast strains used are those of budding yeast, in particular yeasts of the Kluyveromyces type. The K. lactis MW98-8C strain (a. uraA, aro, lys, K⁺, pKD1^o) and K. lactis CBS 293.91 were used in particular. A sample of the

MW98-8C strain was deposited on September 16, 1988 at Centraalbureau voor Schimmelkulturen (CBS), Baarn (Netherlands) where it was registered under the number CBS 579.88.

15.

A bacterial strain (E. coli) transformed with the plasmid pET-8c52K was deposited on April 17, 1990 with the American Type Culture Collection under the number ATCC 68306.

Yeast strains transformed by the expression plasmids coding for the proteins of the present invention are cultivated in Erlenmeyer flasks or 2 l pilot fermenters (SETRIC, France), in rich medium (YPD: 1% yeast extract, 2% Bactopeptone, 2% glucose; or YPL: 1% yeast extract, 2% Bactopeptone, 2% lactose) under constant stirring at 28°C.

EXAMPLE 1:

PROTOCOL FOR STRICT PEPTIDE INSERTION UTILIZING A SINGLE RESTRICTION SITE PRESENT IN THE HSA GENE.

Due to its uniqueness in the HSA gene and the associated vector, the site PvuII, located naturally in the coding sequence, is particularly useful as a cloning site for a biologically active peptide that one wishes to insert in translation phase into the HSA at the level of the 58th residue. In a particular embodiment, it is useful to employ peptides from p residues for which the coding sequence is $[3xN]_p$. In this case, the oligonucleotides synthesized are of type 5'-NN $[3xN]_p$ N-3' and its complementary strand. The ligation of this fragment with the restriction fragment HindIII-HindIII Δ PvuII, corresponding to the totality of the gene coding for the HSA, generates a HindIII-HindIII restriction fragment comprising a hybrid gene coding for a chimeric protein of strict insertion type HSA. In another embodiment, the peptide may be repeated several times in the chimera.

EXAMPLE 2:

TOTAL OR PARTIAL PEPTIDE INSERTION PROTOCOL UTILIZING TWO SINGLE AND NATURAL RESTRICTION SITES.

In a particular embodiment, the use of two single restriction sites allows for the creation of genes coding for chimeras having the active peptide in strict insertion or partial insertion.

For example, the existence of single HincII and AvrII sites in the coding sequence of HSA and in the vector allows an HindIII-HindIII Δ HincII-AvrII fragment to be generated. The elimination of the nucleotide fragment HincII-AvrII corresponds to the deletion of the peptide fragment T(420) – N(429). The use of an appropriate complementary oligonucleotide permits a peptide in cloning phase to be cloned in the HSA gene. This restriction fragment may be exclusively the complementary coding sequence of the peptide, or may correspond to a blend of the coding sequence of the active peptide and the coding sequence of the T(420) – N(429) fragment of the HSA. The active peptide may be present several times in the chimera.

EXAMPLE 3 EXPRESSION PLASMIDS

The chimeric proteins of the preceding examples may be expressed in the yeast, starting from functional promoters, regulated or constitutive, such as, for example, those present in the plasmids pYG105 (promoter LAC4 of Kluyveromyces lactis), pYG106 (promoter PGK of Saccharomyces cerevisiae), pYG536 (promoter PHO5 of S. cerevisiae), or the hybrid promoters such as those described in the patent application EP 361 991. The plasmids pYG105 and pYG106 are particularly useful here because they permit the expression of genes coded by the restriction fragments HindIII that are described in the preceding examples and are cloned in the HindIII site and in productive orientation (defined as the orientation that places the "prepro" region of the albumin in proximal fashion relative to the transcription promoter), from the functional promoters in K. lactis, whether regulatable (pYG105), or constituent (pYG106). The plasmid pYG105 corresponds to the plasmid pKan707 described in patent application EP 361 991, in which the restriction site HindIII, single and located in the gene for resistance to geneticine (G418), has been destroyed by directed mutagenesis while preserving a protein unchanged (oligodeoxynucleotide 5'-GAAATGCATAAGCTCTTGCCATTCTCACCG-3' (SEQ ID No. 3). The fragment SaII-SacI coding for the gene URA3 of the transferred plasmid was replaced by a restriction fragment SaII-SacI, comprising an expression cassette consisting of the promoter LAC4 of K. lactis (in the form of a fragment SaII-HindIII) and the terminator of the PGK gene of S. cerevisiae (in the form of a fragment HindIII-SacI). The plasmid pYG105 is very stable mitotically in the Kluyveromyces yeasts. It

17.

is represented in Figure 3. The plasmids pYG105 and pYG106 do not differ from each other than in the nature of the transcription promoter encoded by the fragment SaII-HindIII.

EXAMPLE 4:

TRANSFORMATION OF YEASTS

The transformation of yeasts belonging to the Kluyveromyces type, and particularly the strains MW98-8C and CBS 293.91 of K. lactis, is carried out, for example, by the technique of treating the intact cells with lithium acetate [Ito H. et al., J. Bacteriol. 153 (1983) 163-168], adapted as follows. The growth of the cells takes place at 28°C in 50 ml of YPD medium through stirring and up to an optical density between 0.6 and 0.8 at 600 nm (OD₆₀₀); the cells are harvested by centrifugation at low speed, washed in a sterile solution of TE (10 mM Tris HCl pH 7.4; 1 mM EDTA), resuspended in 3-4 ml of lithium acetate (0.1 M in the TE) to obtain a cellular density of about 2×10^8 cells/ml, then incubated at 30°C for 1 hour with moderate stirring. Aliquots of 0.1 ml of the resulting suspension of competent cells are incubated at 30°C for 1 hour in the presence of ADN and at a final concentration of 35% polyethylene glycol (PEG₄₀₀₀, Sigma). After a thermal shock of 5 minutes at 42°C, the cells are washed twice, resuspended in 0.2 ml of sterile water and incubated for 16 hours at 28°C in 2 ml of YPD medium to permit the phenotypic expression of the gene conferring resistance to G418, expressed under control of the promoter P_{kl} (cf. EP 361 991); 200 µl of the cellular suspension is then spread over YPD selective containers (G418, 200µg/ml). The containers are put into incubation at 28°C and the transformed cells appear after 2 to 3 hours of cellular growth.

EXAMPLE 5:

SECRETION OF CHIMERAS

After selection in rich medium supplemented with G418, the recombinant clones are tested for their capacity to secrete the mature form of chimeric proteins. Clones corresponding to the CBS 293.91 or MW98-8C strain, transformed by the expression plasmids of the chimeras between the HSA and the biologically active portion are incubated in YPD or YPL medium at 28°C. Cells in the supernatant are retrieved by centrifugation when the cells reach stationary growth phase, concentrated tenfold if necessary by precipitation for 30 minutes at 20°C in a final concentration of 60% ethanol, and tested after electrophoresis on SDS-PAGE 0 gel, either directly by coloration of the gel by coomassie blue, or after

immunoblot, using primary antibodies directed against the biologically active portion or a polyclonal rabbit serum directed against the HSA. During the immunologic detection experiments, the nitrocellulose filter is first incubated in the presence of the specific primary antibody, washed several times, incubated in the presence of goat antibodies directed against the primary antibody, then incubated in the presence of an avidine-peroxidase complex utilizing the "ABC kit" distributed by Vectastain (Biosys S.A., Compiègne, France). The immunologic reaction is then revealed by addition of diamino-3,3' benzidine tetrachlorhydrate (Prolabo) in the presence of oxygenated water, according to the manufacturer's recommendations.

EXAMPLE 6: PURIFICATION OF CHIMERAS

The chimeras present in the supernatants of the cultures corresponding to the transformed strain CBS 293.91 are characterized at a first stage with the aid of antibodies specific to the HSA part. It may be desirable to purify some of these chimeras. The culture is then centrifuged (10,000 g, 30 min), the supernatant is passed through a 0.22 mm filter (Millipore), then concentrated by ultrafiltration (Amicon) using a membrane for which the discrimination threshold is at 30 kDa. The concentrate obtained is then dialyzed using a solution of Tris HCl (50 mM, pH 8) and purified by column. For example, the concentrate corresponding to the supernatant of the culture from the transformed strain CBS 293.91 is purified by affinity chromatography on Trisacryl Blue (IBF), and the samples are next dialyzed using water. Purification by molecular sieve may then be carried out. In this case, a Superose 12 column (Pharmacia) is equilibrated beforehand with a buffer of 20 mM NaH_2PO_4 and 100 mM NaCl at pH 7.0, and the samples resulting from the affinity chromatography are loaded on the column, harvested and characterized. Purification by ion exchange chromatography may also be used. In certain cases, the concentrate obtained after ultrafiltration is dialyzed using a solution of Tris HCl (50 mM, pH 8), then deposited in 20 ml fractions on a cation exchange column (5 ml) (S Fast Flow, Pharmacia) equilibrated with the same buffer. The column is then washed several times with the Tris HCl solution (50 mM, pH 8), and the chimera protein is then eluted in the column by an NaCl gradient (0 to 1 M).

20.

The fractions containing the chimera protein are then combined and dialyzed using a solution of Tris HCl 50mM (pH 8) and redeposited on the S Fast Flow column. After elution of the column, the fractions containing the protein are combined, dialyzed using water and lyophilized before characterization.

EXAMPLE 7:

INSERTION OF PEPTIDE 11 IN THE ALBUMIN SEQUENCE

Peptide 11 was described as an epitope of tryptophan synthase (Larvor et al., Mol. Immunol. (1991) 28, 523-531).

1. Strict insertion in a region naturally having single restriction sites

Peptide 11 has the following peptide sequence in the terminal N a C sense: HGRVGIYFGMK (SEQ ID No.20). The strategy of insertion in the HincII-AvrII zone has consisted of creating a ligation between a synthesized nucleotide fragment coding for peptide 11 such that the reading frame for the sequence coding for HSA is respected, and the nature of the sequence coding for HSA is as it originally was. In a first stage we have thus synthesized the following two oligonucleotides: 5'-CCATGGTAGAGTAGGTATCTATTTTCGGTATGAAAACCTCCAACCTTG TAGAGGTCTCGAGAAAT-3' (SEQ ID No.4) and 5'-CTAGATTTCTCGAGACCTCTACAAGATGTGGAGTTTTTCATACCGA AATAGATACCTACTCTACCATGG-3' (SEQ ID No.5).

These two oligonucleotides have been hybridized and then ligated from the fragment HindIII-HindIII Δ HincII-AvrII, thus generating the totality of the HSA gene comprising the total insertion of peptide 11 between S(419) and T(420), immediately preceding the "prepro" exportation region of the HSA. This fragment is cloned in productive orientation and in the HindIII site of the pYG105 plasmid.

21.

2. Insertion by substitution and addition in a region naturally having single restriction sites

In another embodiment, the oligonucleotides synthesized were the following: 5'-CCATGGTAGAGTAGGTATCTATTTCGGTATGAAA-3' (SEQ ID No.6), and 5'-CTAGTTTCATACCGAAATAGATACCTACTTCTTACCATGG-3' (SEQ ID No.7). These two oligonucleotides are hybridized and then ligated to the fragment HindIII-HindIII Δ HincII-AvrII, thus generating the gene of a chimera severed from the residue T(420) to N(429) and substituted by the sequence of peptide 11. This fragment is cloned in productive orientation and in the HindIII site of the pYG105 plasmid.

3. Strict insertion in a region not having single restriction sites

In another embodiment, two single restriction sites, Sst I and Xho 1, have been created by directed mutagenesis. The same type of directional cloning was achieved. In the case of total insertion of peptide 11 between residue A(191) and S(192), the following two oligonucleotides were synthesized:

5'-ACGGGATGAAGGGAAGGCCCATGGTAGAGTAGGTATCTATTT
CGGTATGAAA-3' (SEQ ID No.8), and

5'-TCGATTTCATTACCGAAATAGATACCTACTCTACCATGGGCCTT
CCCTTCATCCCGTAGCT-3' (SEQ ID No.9).

One then proceeds according to the protocols already described in the preceding points 1 or 2 to obtain the corresponding expression plasmid.

4. Insertion by substitution in a region not having single restriction sites

The insertion requires the preliminary creation of two restriction sites, SstI and XhoI. For the partial insertion itself, the following two oligonucleotides were synthesized:

5'-ACATGGTAGAGTAGGTATCTATTTCGGTATGAAA-3'
(SEQ ID No.10), and

5'-TCGATTTCATACCGAAATAGATACCTACTCTACCATGTAGCT-3'
(SEQ ID No.11).

In this last case, the residues R(186) to A(191) are deleted and substituted by the peptide 11. The two expression plasmids of these chimeras are 1671 and 1667, respectively.

Each of the plasmids obtained in points 1, 2, 3 and 4 is used to transform a yeast strain, according to the protocol described in example 4. The corresponding proteins are secreted and purified as in examples 5 and 6.

EXAMPLE 8:

INSERTION OF PEPTIDE SEQUENCE IEGR, THE SUBSTRATE OF FACTOR Xa

The insertion of the peptide containing the sequence IEGR (SEQ ID No.21) comes next, target of the protease factor Xa, which turns prothrombin into thrombin in the cascade of reactions involved in the coagulation of blood.

1. Strict insertion in region 5 having a single PvuII site.

To achieve this insertion at the PvuII site of the albumin gene, a first step is to create a replicative vector lacking a PvuII site, in which the gene coding for the prepro albumin is then inserted. The following two complementary oligonucleotides, coding for the sequence IEGR, have also been synthesized: 5'- GATCCATAGAAGGTCGACTAG-3' (SEQ ID No.12), and 3'- CTAGGTATCTTCCAGCTGATC-5' (SEQ ID No.13).

These two oligonucleotides are then hybridized and inserted at the level of site PvuII in the albumin gene. The modifications in the nucleotide and peptide sequences of the albumin that their insertion causes to appear are presented in Figure 4. The cloning strategy is schematized in Figure 5.

This construction illustrates the case where the junction sequences are introduced on either side of the peptide at issue.

2. Insertion in region 13.

Figure 6 depicts a cloning strategy for a peptide in region 13.

In what follows, the restriction sites Sst I and Xho 1 were created by directed mutagenesis. In the case of region 13, a total substitution and an addition substitution of the IEGR sequence were carried out respectively, either simple or surrounded by the junction sequences.

23.

In regard to the total substitution, the oligonucleotides utilized are the following:

5'-CAGAATCGAAGGTAGAGCC-3' (SEQ ID No.14), and

5'-TCGAGGCTCTACCTTCGATCGAGGGTAGCT-3' (SEQ ID NO.15).

On the protein level, the sequence (187)DEGK (SEQ ID No.22) is substituted by IEGR.

In the second case, the oligonucleotides utilized are the following:

5'-ACCCTCGATCGAAGGTAGATCTCCA-3' (SEQ ID No.16)

5'-TCGATGGAGATCTACCTTCGATCGAGGGTAGCT-3'

(SEQ ID No.17)

On the protein level, the albumin is severed from residue R(186) to residue A(191) and replaced by the sequence PSIEGRSP (SEQ ID No.23), thus leading to an addition of two residues.

The corresponding proteins are secreted and purified according to the protocols described in the preceding examples. Table 1 below lists the characteristics of these chimeras.

Table 1

Albumin Incorporating	Expression Rate μg/l	Purification Yield	Final Purity
IEGR	150	15	98
PSIEGRPS	110	36	96

3. Biological activity of the chimera obtained according to example 8.1

The chimera is incubated for 3 hours at 37°C in the presence of bovine factor Xa in an enzyme/ substrate ratio of 1/10 and a buffer of Tris 50mM, NaCl 100mM, and CaCl₂ 1mM, at pH 8.0. At the conclusion of this treatment, an SDS PAGE analysis is performed that provides evidence for cleavage of the chimera, characterized by the generation of an albumin fragment with a mass of the order of 80 kD.

LIST OF SEQUENCES

(1) GENERAL INFORMATION:

- (i) APPLICANT:
 - (A) NAME: RHONE-POULENC RORER S. A.
 - (B) STREET: 20, avenue Raymond ARON
 - (C) CITY: ANTONY
 - (D) COUNTRY: FRANCE
 - (E) ZIP CODE: 92165
- (ii) TITLE OF THE INVENTION: New biologically active polypeptides, their preparation and pharmaceutical compositions containing them.
- (iii) NUMBER OF SEQUENCES: 24
- (iv) LEGIBLE FORM FOR COMPUTER:
 - (A) TYPE OF SUPPORT: TAPE
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (OEB)

(2) INFORMATION FOR SEQ ID NO.1:

- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 13 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (xi) DESCRIPTION OF SEQUENCE: SEQ ID NO.1

GGCCNNNNNG GCC

(3) INFORMATION FOR SEQ ID NO.2:

- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 8 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO.2

GCGGCCGC

25.

(4) INFORMATION FOR SEQ ID NO.3:

- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO.3

GAAATGCATA AGCTCTTGCC ATTCTCACCG

(5) INFORMATION FOR SEQ ID NO.4:

- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 64 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO.4

CCATGGTAGA GTAGGTATCT ATTTCGGTAT GAAAACTCCA ACTCTTGTAG
AGGTCTCGAG AAAT

(6) INFORMATION FOR SEQ ID NO.5:

- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 68 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO.5

CTAGATTTCT CGAGACCTCT ACAAGATGTG GAGTTTTCAT ACCGAAATAG
ATACCTACTC TACCATGG

26.

(7) INFORMATION FOR SEQ ID NO.6:

- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii)ANTI-SENSE: NO
- (xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.6

CCATGGTAGA GTAGGTATCT ATTCGGTAT GAAA

(8) INFORMATION FOR SEQ ID NO.7:

- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 40 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii)ANTI-SENSE: NO
- (xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.7

CTAGTTTCAT ACCGAAATAG ATACCTACTT CTTACCATGG

(9) INFORMATION FOR SEQ ID NO.8:

- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 52 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii)ANTI-SENSE: NO
- (xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.8

ACGGGATGAA GGGAAGGCC ATGGTAGAGT AGGTATCTAT TTCGGTATGAA
AA

(10) INFORMATION FOR THE SEQ ID NO.9:

- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 61 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: ADNc

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO.9

TCGATTTCAT TACCGAAATA GATACCTACT CTACCATGGG CCTTCCCTTC
ATCCCGTAGC T

(11) INFORMATION FOR SEQ ID NO.10:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 34 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: simple

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: ADNc

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO.10

ACATGGTAGA GTAGGTATCT ATTCGGTAT GAAA

(12) INFORMATION FOR SEQ ID NO.11:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 42 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: simple

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: ADNc

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO.11

TCGATTTCAT ACCGAAATAG ATACCTACTC TACCATGTAG CT

(13) INFORMATION FOR SEQ ID NO.12:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: simple

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: ADNc

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO.12

GATCCATAGA AGGTCGACTA G

(14) INFORMATION FOR SEQ ID NO.13:

(i) CHARACTERISTICS OF THE SEQUENCE:

WO 95/30759

PCT/FR95/00520

(A)LENGTH: 21 base pairs
(B)TYPE: nucleic acid

- (C) NUMBER OF STRANDS: simple
- (E) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO.13

CTAGGTATCT TCCAGCTGAT C

(15) INFORMATION FOR SEQ ID NO. 14:

- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 14

CAGAATCGAA GG TAGAGCC

(16) INFORMATION FOR SEQ ID NO. 15:

- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 15

TCGAGGCTCT ACCTTCGATC GAGGGTAGCT

17) INFORMATION FOR SEQ ID NO. 16:

- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 16

ACCCTCGATC GAAGGTAGAT CTCCA

(18) INFORMATION FOR SEQ ID NO. 17:

(i) CHARACTERISTICS OF THE SEQUENCE:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) NUMBER OF STRANDS: simple
- (D) CONFIGURATION: linear
- (ii) TYPE OF MOLECULE: ADNc
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 17

TCGATGGAGA TCTACCTTCG ATCGAGGGTA GCT

- 19) INFORMATION FOR SEQ ID NO. 18:
- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
 - (ii) TYPE OF MOLECULE: ADNc
 - (iii) HYPOTHETICAL: NO
 - (iii) ANTI-SENSE: NO
 - (xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 18

GTCCCGGATG GAGCGCGTAC TTAGAGAGAA T

- 20) INFORMATION FOR SEQ ID NO. 19:
- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 39 base pairs
 - (B) TYPE: nucleic acid
 - (C) NUMBER OF STRANDS: simple
 - (D) CONFIGURATION: linear
 - (ii) TYPE OF MOLECULE: ADNc
 - (iii) HYPOTHETICAL: NO
 - (iii) ANTI-SENSE: NO
 - (xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 19

TCGACAGGGC CTACCTCGCG CATGAATCTC TCTTAAGCT

- 21) INFORMATION FOR SEQ ID NO. 20:
- (i) CHARACTERISTICS OF THE SEQUENCE:
 - (A) LENGTH: 11 amino acids
 - (B) TYPE: amino acid
 - (C) CONFIGURATION: linear
 - (xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 20

His Gly Arg Val Gly Ile Tyr Phe Gly Met Lys

22) INFORMATION FOR SEQ ID NO. 21:

(i) CHARACTERISTICS OF THE SEQUENCE:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) CONFIGURATION: linear

(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 21

Ile Glu Gly Arg

23) INFORMATION FOR SEQ ID NO. 22:

(i) CHARACTERISTICS OF THE SEQUENCE:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) CONFIGURATION: linear

(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 22

Asp Glu Gly Lys

24) INFORMATION FOR SEQ ID NO. 23:

(i) CHARACTERISTICS OF THE SEQUENCE:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (C) CONFIGURATION: linear

(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 23

Pro Ser Ile Glu Gly Arg Ser Pro

25) INFORMATION FOR SEQ ID NO. 24:

(i) CHARACTERISTICS OF THE SEQUENCE:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (C) CONFIGURATION: linear

(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 24

Arg Met Glu Arg Val Leu Arg Glu Asn

CLAIMS

1. Recombinant polypeptide comprising at least one active portion derived from a polypeptide, natural or synthetic, biologically active, genetically inserted in an albumin or one of its variants or derivatives.
2. Polypeptide according to claim 1, characterized in that the biologically active peptide has therapeutic activity and is of human origin.
3. Polypeptide according to claim 2, characterized in that the polypeptide having a therapeutic activity is chosen from among all or part of the enzymes, enzyme inhibitors, antigens, antibodies, hormones, receptors, coagulation factors, interferons, cytokines, growth and/or differentiation factors, factors involved in the genesis/resorption of bone tissue, chemotactic factors, motility or cellular migration factors, cytostatic factors, bactericidal or antifungal factors, or plasmatic, interstitial or extracellular matrix adhesive molecules.
4. Polypeptide according to one of claims 1 through 3, characterized in that the polypeptide having a therapeutic activity is chosen from among any antagonist or agonist peptide sequence having molecular and/or cellular interaction involved in the pathologies of circulatory and interstitial compartments.
5. Polypeptide according to one of claims 1 through 4, characterized in that the active portion has a structure chosen from among:
 - a) the entire peptide structure or,
 - b) a fragment of (a) or a structure derived from (a) by structural modification (mutation, substitution, addition and/or deletion of one or more residues) and preserving therapeutic activity.
6. Polypeptide according to one of claims 1 through 5, characterized in that the active portion is inserted strictly into the interior of the albumin or surrounded by junction sequences.

7. Polypeptide according to one of claims 1 through 6, characterized in that the active portion is inserted preferably at the level of those regions on the albumin that presumably form regions exposed to the surface of the molecule.
8. Polypeptide according to one of claims 1 through 7, characterized in that the active portion is inserted at the level of region 5 extending from residues 57 to 62 of the albumin.
9. Polypeptide according to one of claims 1 through 7, characterized in that the active portion is inserted at the level of region 8 extending from residues 103 to 120 of the albumin.
10. Polypeptide according to one of claims 1 through 7, characterized in that the active portion is inserted at the level of region 13 extending from residues 178 to 200 of the albumin.
11. Polypeptide according to one of claims 1 through 7, characterized in that the active portion is inserted at the level of the region from residue 415 to residue 425 delimited by the helices h2 and h3 of domain III in the albumin.
12. Polypeptide according to one of claims 1 through 12, characterized in that the active portion is inserted in a single or multiple manner.
13. Polypeptide according to claim 12 characterized in that the active portion is repeated several times in the same place and/or in different regions of the albumin.
14. Polypeptide according to one of claims 1 through 13, characterized in that the active portions inserted are different in nature.
15. Variant of a nucleotide sequence coding for the albumin or one of its variants or derivatives including at least one single non-natural restriction site.
16. Nucleotide sequence coding for a polypeptide according to any of claims 1 to 14.
17. Nucleotide sequence according to claim 16 characterized in that it comprises a "leader" sequence allowing the secretion of the expressed polypeptide.

18. Expression cassette comprising a nucleotide sequence according to claim 16 or 17 under the control of a transcription initiation region and possibly a transcription termination region.
19. Self-replicating plasmid comprising an expression cassette according to claim 18.
20. Eukaryote or prokaryote recombinant cell in which a nucleotide sequence according to claim 16 or 17 or an expression cassette according to claim 18 or a plasmid according to claim 19 has been inserted.
21. Recombinant cell according to claim 20 characterized in that it is a yeast, an animal cell, a fungus or a bacterium.
22. Recombinant cell according to claim 21 characterized in that it is a yeast.
23. Recombinant cell according to claim 22 characterized in that it is a yeast of the type Saccharomyces or Kluyveromyces.
24. Process for preparing a polypeptide as defined in one of claims 1 to 14 characterized in that a recombinant cell is cultivated according to one of claims 20 to 23 under expression conditions, and the polypeptide produced is recovered.
25. Pharmaceutical composition comprising one or more peptides according to any of claims 1 to 14.
26. Pharmaceutical composition comprising a nucleotide sequence according to either claim 16 or 17 that may be used in gene therapy.

WO 93/30751

PCT/FR95/00520

1/6

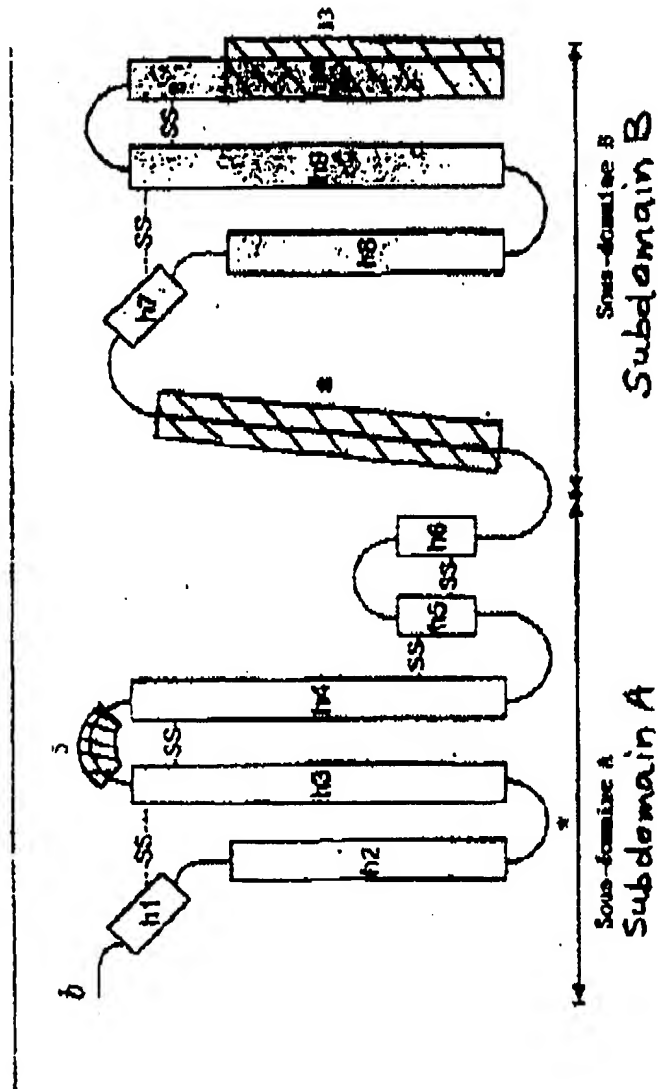


Figure 1

WO 95/30759

PCT/FR93/00520

3/6

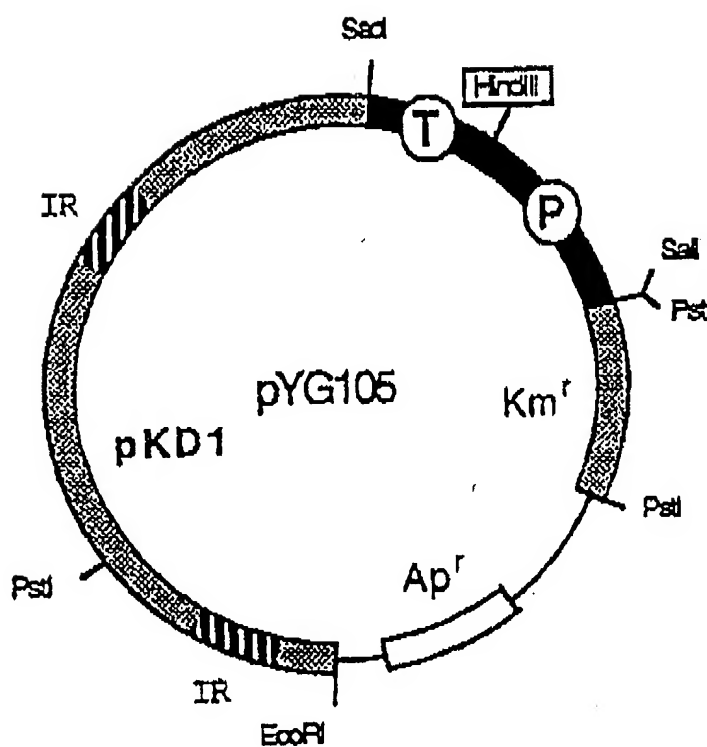


Figure 3

PCT/FR93/00520

a

GAG TCA GGA TCC ATA GAA GGT CGA CTA GCT GAA AAT
E S G S I E G R L A E N

b

GAG TCA G CT GAA AAT

Pvu II

Figure 4

WD 9530754

PC107094-00000

5/6

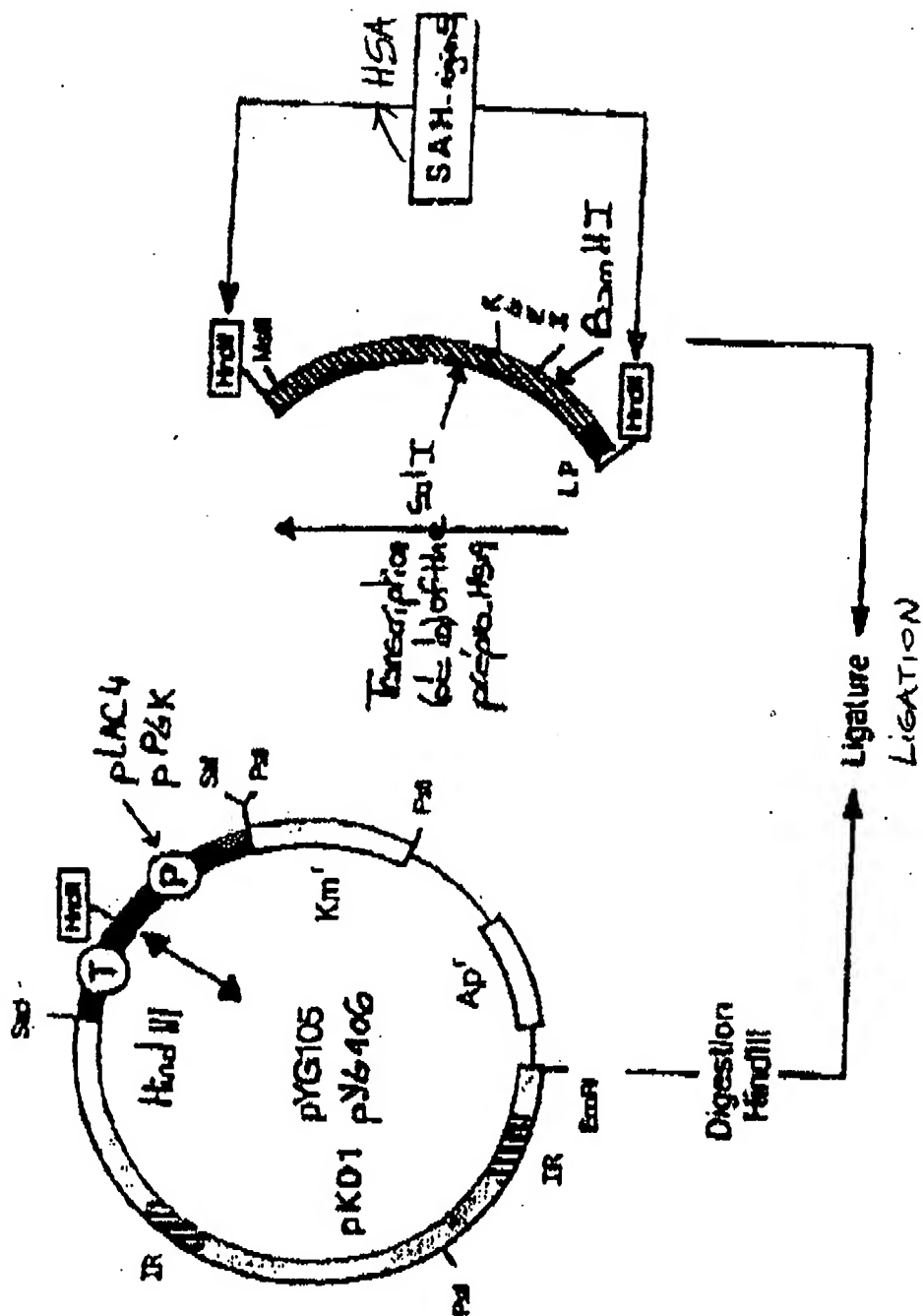


Figure 3

WQ 95/30759

PCT/FR95/0630

6/6

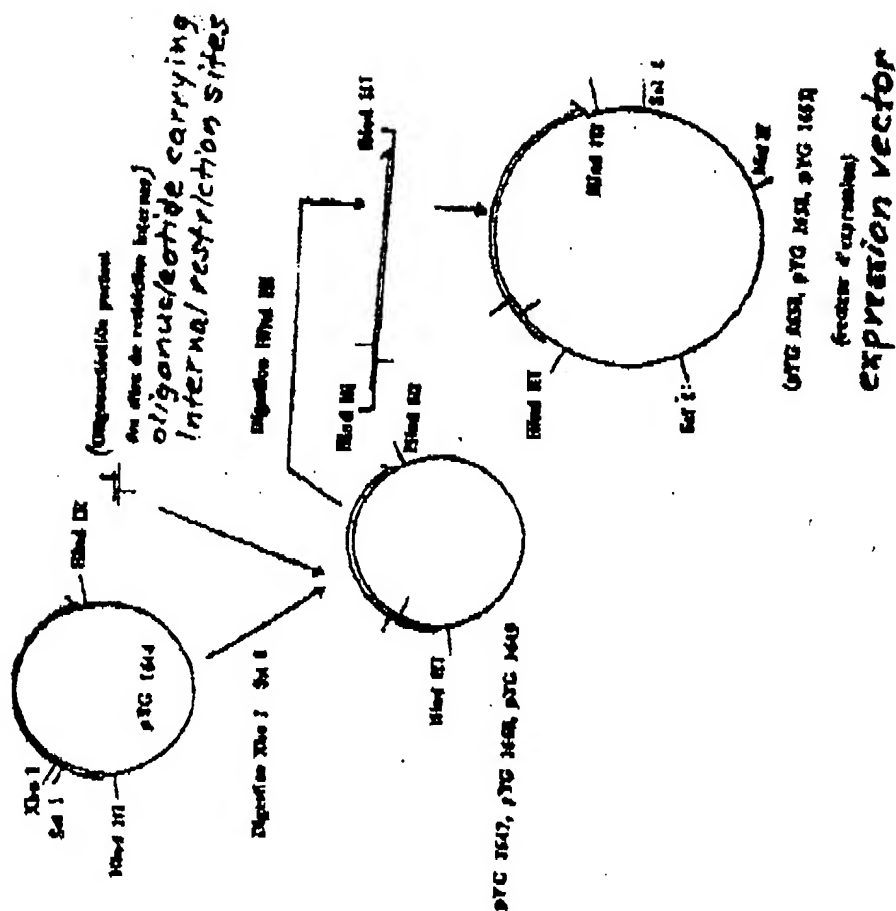


Figure 6

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.